

From: Ford, Vanessa
S nt: Thursday, October 11, 2001 1:28 PM
To: STIC-Biotech/ChemLib
Subject: In re: 09596101 Sequence Search

Please search SEQ ID NO: 1.
Please search SEQ ID NO. 3.
Please include an interference search.

Vanessa L. Ford
Biotechnology Patent Examiner
Office: CM1 8D17 ✓
Mailbox: CM1 8E12 ✓
Phone: 703.308.4735

Art unit Please

RECEIVED
OCT 11 2001
STIC

P

SEARCH REQUEST FORM

Access DB#

52795

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

Searcher: Al Nasser
Searcher Phone #: 308-4581
Searcher Location: Biotec lib.
Date Searcher Picked Up: 10/11/01
Date Completed: 10/11/01
Searcher Prep & Review Time: _____
Clinal Prep Time: 6 min
Online Time: 2 min

Type of Search

NA Sequence (#) _____
AA Sequence (#) 2
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbi _____
Dr Link _____
Lexis/Nexis _____
Sequence Systems AB5502
WWW/Internet _____
Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:48:17 ; Search time 35.17 Seconds
(without alignments)
22.409 Million cell updates/sec

Title: US-09-596-101A-1
Perfect score: 64
Sequence: 1 SGEIDITFTGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
al number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_0601.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	13	AA124915	Eisenia foetida co
2	64	100.0	20	AA124914	Eisenia foetida co
3	50	78.1	306	AA197362	Oerskovia beta-1.3
4	45	70.3	294	AA152463	Mycobacterium tube
5	44	68.8	263	AA129455	Oerskovia xanthine
6	44	68.8	303	AA129457	Oerskovia xanthine
7	44	68.8	435	AA129456	Oerskovia xanthine
8	39	60.9	261	AA11599	Beta-1,3-glucanase
9	39	60.9	232	AA188406	Trichoderma harziz
10	38	59.4	223	AA12318	Corn beta-carotene
11	37	57.8	275	AA14125	Arabidopsis thalia

12	37	57.8	324	21	AA14124
13	37	57.8	336	21	AA14123
14	36	56.2	222	22	AA100115
15	36	56.2	419	22	AA10788
16	36	56.2	1149	22	AA197640
17	36	56.2	1194	20	AA191071
18	36	56.2	1194	21	AA19432
19	36	56.2	1194	22	AA197649
20	36	56.2	1205	20	AA191072
21	36	56.2	1205	22	AA197637
22	36	56.2	1237	22	AA197638
23	36	56.2	1248	22	AA197636
24	36	56.2	1248	22	AA197639
25	36	56.2	1248	22	AA197641
26	36	56.2	1248	22	AA197642
27	36	56.2	1248	22	AA197643
28	36	56.2	1248	22	AA197644
29	36	56.2	1248	22	AA197645
30	36	56.2	1248	22	AA197646
31	36	56.2	1248	22	AA197647
32	36	54.7	175	19	AA186077
33	35	54.7	176	19	AA185821
34	35	54.7	408	21	AA136267
35	35	54.7	430	18	AA134687
36	35	54.7	430	18	AA124482
37	35	54.7	430	19	AA171618
38	35	54.7	438	21	AA136266
39	35	54.7	485	21	AA136265
40	35	54.7	653	21	AA154211
41	35	54.7	690	21	AA170475
42	35	54.7	718	21	AA151607
43	35	54.7	783	21	AA154210
44	35	54.7	783	21	AA154209
45	35	54.7	1196	20	AA132017

ALIGNMENTS

RESULT 1	
AA124915	standard; peptide: 13 AA.
ID	AA124915
XX	AA124915
AC	AA124915
XX	AA124915
DT	25-AUG-1999 (first entry)
XX	25-AUG-1999
DE	Eisenia foetida coelomic cytolytic factor 1 peptide.
XX	Eisenia foetida; coelomic cytolytic factor 1; CCF-1; cancer;
KW	trypanosomal infection; bacterial infection; tumour therapy;
KW	Inflammation; immunology.
XX	
OS	Eisenia foetida.
XX	
PN	WO9931229-A2.
XX	
PD	24-JUN-1999.
XX	
PF	16-DEC-1998; 98MO-EP08169.
XX	
PR	17-DEC-1997; 97EP-0203974.
XX	
PA	(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX	
PI	De Baetselier P;
XX	
DR	WPI; 1999-385905/32.
XX	
PT	Eisenia foetida polypeptides derived from coelomic cytolytic factor 1
XX	
PS	Claim 1; Page 45; 49pp; English.
XX	

Arabidopsis thalia
Arabidopsis thalia
Bacillus lichenifo
N. magadali bacter
Apaf-1Xt/Delta2-10
Apoptosis inducer
Human full-length
Wild type Apaf-1 p
Apoptosis inducer
Apaf-1L protein se
Apaf-1L protein s
Apaf-1Xt protein s
Apaf-1Xt-M368L pro
Apaf-1Xt-L10A prot
Apaf-1Xt-D27A prot
Apaf-1Xt-E39Q prot
Apaf-1Xt-K63Q prot
Apaf-1Xt-L83A prot
Apaf-1Xt-K160R prot
Apaf-1Xt-M368L/K16
S. pneumoniae deri
S. pneumoniae deri
Arabidopsis thalia
Class II EPSP synt
Class II EPSPS use
Staphylococcus aur
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human cyclic nucle
Human soluble guan
Arabidopsis thalia
Arabidopsis thalia
Mouse cation chann

CC The present sequence represents a *Eisenia foetida* coelomic cytolytic
CC factor 1 (CCF-1) peptide. The CCF-1 protein has antiparasitic,
CC antibacterial and anti-inflammatory activity. Recombinant coelomic
CC cytolytic factor 1 (rCCF-1) is trypanolytic for the African trypanosome
CC *Trypanosoma brucei* in a dose-dependent manner. The trypanolytic activity
CC of rCCF-1 can be inhibited by anti-CCF-1 and anti-tumour necrosis factor
CC (TNF)/rTP monoclonal antibodies. Furthermore, N,N'-diacetylchitobiose
CC inhibits potentially trypanolytic activity of rCCF-1. These data
CC corroborate the findings that CCF-1 shares a trypanolytic, lectin-like
CC domain with TNF-alpha. CCF-1 is useful to treat trypanosomal or
CC bacterial infections or cancer. The proteins and peptides are also
CC useful in tumour therapy. Inflammation and other areas of immunology.
CC The annelid peptide is derived from a 42 kDa cytolytic protein named
CC CCF-1 that binds lipopolysaccharide and beta-1,3-glucan. The factor
CC resembles the vertebrate tumour necrosis factor-alpha (TNF-alpha), and
CC may be used as an alternative for TNF-alpha.

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
|||||
Db 1 sgeidlietignr 13

RESULT 2

AAV24914
ID AAV24914 standard; Protein; 384 AA.

AC AAV24914;

XX 25-AUG-1999 (first entry)

DE *Eisenia foetida* coelomic cytolytic factor 1 protein.XX *Eisenia foetida*; coelomic cytolytic factor 1; CCF-1, cancer;

KM trypanosomal infection; bacterial infection; tumour therapy;

XX inflammation; immunology.

OS *Eisenia foetida*.

FH Key Location/Qualifiers

FT Peptide 1..17
/label= signal
FT 18..384
/label= CCF-1

XX Protein

XX WO9931229-A2.

XX 24-JUN-1999.

XX 16-DEC-1998; 98WO-EP08169.

XX 17-DEC-1997; 97EP-0203974.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOL.

XX De Baetseller P;

XX WPI; 1999-385905/32.

XX N-PSDB; AAX83611.

XX *Eisenia foetida* polypeptides derived from coelomic cytolytic factor 1

XX Claim 2; Page 48-49; 49pp; English.

XX The present sequence represents the *Eisenia foetida* coelomic cytolytic
CC factor 1 (CCF-1). The protein has antiparasitic, antibacterial and
CC anti-inflammatory activity. Recombinant coelomic cytolytic factor 1

CC (rCCF-1) is trypanolytic for the African trypanosome *Trypanosoma brucei*
CC in a dose-dependent manner. The trypanolytic activity of rCCF-1 can be
CC inhibited by anti-CCF-1 and anti-tumour necrosis factor (TNF)/rTP
CC monoclonal antibodies. Furthermore, N,N'-diacetylchitobiose inhibits
CC potentially trypanolytic activity of rCCF-1. These data corroborate the
CC findings that CCF-1 shares a trypanolytic, lectin-like domain with
CC TNF-alpha. CCF-1 is useful to treat trypanosomal or bacterial infections
CC or cancer. The proteins and peptides are also useful in tumour therapy,
CC inflammation and other areas of immunology. The annelid peptide is
CC derived from a 42 kDa cytolytic protein named CCF-1 that binds
CC lipopolysaccharide and beta-1,3-glucan. The factor resembles the
CC vertebrate tumour necrosis factor-alpha (TNF-alpha), and may be used as
CC an alternative for TNF-alpha.

SQ Sequence 384 AA;

Query Match 100.0%; Score 64; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
|||||
Db 178 sgeidlietignr 190

RESULT 3

AAR97362
ID AAR97362 standard; Protein; 306 AA.

AC AAR97362;

XX 03-OCT-1996 (first entry)

DE *Oerskovia beta*-1,3-glucanase.XX *Beta*-1,3-glucanase; *Cellulomonas cellulans*; *Bacillus subtilis*;KM lytic enzyme; *Beta*-glucan degradation; cell wall lysis;

XX pigment; colorant; flavour; yeast extract; protoplast.

XX *Oerskovia xanthoneolytica* strain LG109 (DSM 10297).

OS

FH Key Location/Qualifiers

FT Peptide 1..35
/label= sig_peptide
FT 36..63
/label= Pro-peptide
FT 64..306
/label= Mat_protein

XX Protein

XX WO9612013-A1.

XX 25-APR-1996.

XX 16-OCT-1995; 95WO-DK00414.

XX 14-OCT-1994; 94DK-0001192.

XX (NOVO) NOVO-NORDISK AS.

XX Asejo JA; Diers I, Ferrer P, Halkier T, Hedegaard L;

XX Savva D;

XX WPI; 1996-222000/22.

XX N-PSDB; AAT29043.

XX DNA construct encoding enzyme with *Beta*-1,3-glucanase activity -
XX useful for modifying or degrading *Beta*-glucan contg. material and in
XX the prepn. of e.g. food colourants, flavourings and yeast extractsXX Claim 1; Page 42-43; 60pp; English.
XX A novel *Beta*-1,3-glucanase (AAR97362) from *Oerskovia xanthoneolytica*

CC LIG109 is useful for degrading or modifying beta-glucan-contg.
 CC material. Its amino acid sequence was deduced from a genomic
 CC DNA sequence (AAT29043) isolated from an O. xanthineolytica library.
 CC Recombinant beta-1,3-glucanase can be produced on a large scale
 CC using transformed host cells, esp. Bacillus subtilis DN1885 or
 CC TOC46. Protease-free beta-1,3-glucanase can be obtd. that is useful
 CC for lysing fungal cell walls, allowing recovery of intracellular
 CC proteins. The enzyme is also useful for the prepn. of protoplasts
 CC and for the prodn. of pigments, colorants, flavours, yeast extract
 CC and pharmaceuticals.

SO Sequence 306 AA;

Query Match 78.1%; Score 50; DB 17; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.12;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITETIGN 12
 |||||:|:|
 180 sgeidimenygn 191

RESULT 4

ID AAB52463 standard; protein; 294 AA.

AC AAB52463;

DT 23-FEB-2001 (first entry)

DE Mycobacterium tuberculosis secreted protein #28.

KM Mycobacterium tuberculosis secreted protein; MTSP; vaccine.

OS Mycobacterium tuberculosis.

PN W0200066143-A1.

PD 09-NOV-2000.

PF 04-MAY-2000; 2000MO-US12197.

PR 04-MAY-1999; 9905-0132479.

PR 04-MAY-1999; 9905-0132503.

PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.

Gennaro ML, Gomez MJ;

WPI; 2001-007151/01.

Novel Mycobacterium tuberculosis secreted polypeptides and
 tuberculosin -

claim 11; Fig 1; 60pp; English.

The present invention relates to Mycobacterium tuberculosis secreted
 CC proteins (MTPs), where the polypeptide has M. tuberculosis specific
 CC antigenic and immunogenic properties. Compositions of the invention may
 CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
 CC vaccine against M. tuberculosis infection.

SO Sequence 294 AA;

Query Match 70.3%; Score 45; DB 22; Length 294;
 Best Local Similarity 75.0%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDITETIGN 12
 |||||:|:|

Db 169 sgeidlienygn 180

RESULT 5

ID AAM29455 standard; Protein; 263 AA.

AC AAM29455;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica mature beta-1,3-glucanase.

KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LIG109 (DSM 10297).

Key Location/Qualifiers

FT CDS 23..955

FT sig_peptide 23..120

FT mat_peptide 164..952

FT /*tag= c

W09739114-A1.

PD 23-OCT-1997.

PF 14-APR-1997; 97WO-DK00160.

PR 23-AUG-1996; 96DK-0000885.

PR 12-APR-1996; 96DK-0000427.

PA (NOVO) NOVO-NORDISK AS.

PI Diers I, Ferrer P, Halkier T, Hedegaard L;

WPI; 1997-526451/48.

DR N-PSDB; NAT89155.

Example 2; Page 35-36; 64pp; English.

This polypeptide comprises a novel Oerskovia xanthineolytica (OX)

CC enzyme that exhibits beta-1,3-glucanase (BG) activity. Its amino

CC acid sequence was deduced from an isolated genomic DNA sequence

CC (see AAT89155). Claimed DNA constructs that encode the novel BG (see

CC also AAM29456 for corrected sequence), a mannose binding domain (see

CC AAM29458) or a full-length enzyme, i.e. BG with mannose binding

CC domain (see AAM29456), can be used to produce recombinant BG

CC polypeptides, with or without a mannose binding domain, in fungal

CC or bacterial host cells. BG polypeptides are used for the

CC degradation or modification of beta-glucan containing material,

CC especially for the gentle lysis of microbial cell walls, thereby

CC enabling recovery of desirable intracellular products with a

CC reduced amount of contaminants. They can also be used for the

CC production of e.g. pigments, colourants, flavourants, yeast

CC extracts, pharmaceuticals, food or feed compositions, and to

CC prepare protoplasts for use in fusion, transformation and cloning

SO Sequence 263 AA;

Query Match 68.8%; Score 44; DB 18; Length 263;
 Best Local Similarity 72.7%; Pred. No. 1.4;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITFTIG 11
|||||:|:|
Db 116 sgeiditmenvg 126

RESULT 6

AAW29457
ID AAW29457 standard; Protein; 303 AA.
XX AAW29457;
XX
XX
XX 14-APR-1998 (first entry)
XX
XX
XX Oerskovia xanthineolytica beta-1,3-glucanase.
XX
XX
XX Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
KW fungal cell wall; intracellular product; purification; protoplast.
XX
XX Oerskovia xanthineolytica LUG109 (DSM 10297).

Key Location/Qualifiers
FH Peptide 1..52
FT /label= Sig_peptide
FT 53..303
FT Protein /label= Mat_protein

MO9739114-A1.

23-OCT-1997.

14-APR-1997; 97WO-DK00160.

23-AUG-1996; 96DK-0000885.

12-APR-1996; 96DK-0000427.

(NOVO) NOVO-NORDISK AS.

Diers I, Ferrer P, Halkier T, Hedegaard L;

WPI: 1997-526451/48.

N-PSDB; AAT89157.

New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products

Example 2; Page 42-43; 64pp; English.

This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (BG) activity. Its amino acid sequence was deduced from an
CC isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs
CC that encode the novel BG (see also AAW29457), a mannose binding
CC domain (see AAW29458) or a full-length enzyme, i.e. BG with mannose
CC binding domain (see AAW29456), can be used to produce recombinant BG
CC polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. BG polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.

Sequence 303 AA;

Query Match 68.8%; Score 44; DB 18; Length 303;
Best Local Similarity 72.7%; Pred. No. 1.7;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITFTIG 11
|||||:|:|
Db 169 sgeiditmenvg 179

RESULT 7

AAW29456
ID AAW29456 standard; Protein; 435 AA.
XX AAW29456;
XX
XX
XX 14-APR-1998 (first entry)
XX
XX
XX Oerskovia xanthineolytica beta-1,3-glucanase.
XX
XX
XX Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
KW fungal cell wall; intracellular product; purification; protoplast.
XX
XX Oerskovia xanthineolytica LUG109 (DSM 10297).

Key Location/Qualifiers
FH Peptide 1..52
FT /label= Sig_peptide
FT 53..435
FT Protein /label= Mat_protein
FT Domain 304..435
FT /label= Mannose-binding_domain

MO9739114-A1.

23-OCT-1997.

14-APR-1997; 97WO-DK00160.

23-AUG-1996; 96DK-0000885.

12-APR-1996; 96DK-0000427.

(NOVO) NOVO-NORDISK AS.

Diers I, Ferrer P, Halkier T, Hedegaard L;

WPI: 1997-526451/48.

N-PSDB; AAT89156.

New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products

Example 2; Page 39-40; 64pp; English.

This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (BG) activity and which includes a mannose binding domain (MBD).
CC Its amino acid sequence was deduced from an isolated genomic DNA
CC sequence (see AAT89156). Claimed DNA constructs that encode the
CC novel BG lacking a MBD (see AAW29455 and AAW29457), a MBD (see
CC AAW29458), or the full-length enzyme can be used to produce recombinant
CC BG polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. BG polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.

Sequence 435 AA;

Query Match 68.8%; Score 44; DB 18; Length 435;
Best Local Similarity 72.7%; Pred. No. 2.5;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIETIG 11
 |||||:|:|
 Db 169 sgeidmetvny 179

RESULT 8

AAR1599
 ID AAR1599 standard; Protein: 261 AA.
 AC AAR1599;

DT 18-JUN-1991 (first entry)

DE Beta-1,3-glucanase.

KM Alkalophilic; heat resistant.

XX Bacillus sp.

QS JP03053883-A.

XX PD 07-MAR-1991.

XX PF 20-JUL-1989; 89JP-0185928.

XX PR 20-JUL-1989; 89JP-0185928.

XX PA (SHKJ) SHINGIUTSU KAIHATSU.

XX DR WPI; 1991-113290/16.

XX DR N-PSDB; AAQ11293.

PT Heat resistant beta-1,3-glucanase gene DNA - derived from

XX alkali-compatible Bacillus sp.

PS Claim 7; Fig 2; 11pp; Japanese.

CC The new enzyme is derived from alkalophilic Bacillus Sp. The

CC enzyme acts on beta-1,3-glucans to form glucose and laminaribiose.

CC It is stable over a wide pH range with opt. activity at weakly

CC acidic pH and has good heat resistance.

XX Sequence 261 AA;

Query Match 60.9%; Score 39; DB 12; Length 261;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIG 12
 |||||:|:|
 Db 119 geidmetvny 129

RESULT 9

AAR88406
 ID AAR88406 standard; Protein: 292 AA.

XX AAR88406;

DT 19-AUG-1996 (first entry)

DE Trichoderma harzianum endo-1,3(4)-beta-glucanase.

XX Yeast; fungus; enzyme; endo-beta-glucanase; EC-3.2.1.6; hydrolase;

KM endo-1,3(4)-beta-glucanase; beta-glucan degradation; hydrolysis;

XX detergent; surfactant; fungicide; antifungal; cleaning.

OS Trichoderma harzianum CBS 243.71.

XX MO9531533-A1.

XX PD 23-NOV-1995.
 XX 11-MAY-1995; 95WO-DK00188.
 XX 11-MAY-1994; 94DK-0000546.

XX (NOVO) NOVO-NORDISK AS.

XX Andersen LN, Breinholt J, Christgau S, Dalboge H;

PI Kauppinen MS, Kofoed LV, Olsen HS;

XX WPI; 1996-010920/01.

DR N-PSDB; AAT09876.

XX DNA encoding endo-beta-glucanase from Trichoderma harzianum -

PT useful, e.g., in food processing, as antifungal agent, in cleaning

PT compns., etc.

XX Claim 1; Page 40; 57pp; English.

XX This polypeptide is immunologically reactive with antibodies

CC raised against a pure endo-1,3(4)-beta-glucanase from T. harzianum

CC CBS 243.71. Using this sequence, the protein may be expressed

CC recombinantly in transformed host cells, particularly Aspergillus

CC oryzae or Aspergillus niger. Typical applications of the

CC recombinantly produced protein are preparation of protoplasts

CC and yeast extracts, brewing, wine and press-juice manufacture, in

XX foods and feeds, as antifungal agents, etc.

XX Sequence 292 AA;

Query Match 60.9%; Score 39; DB 17; Length 292;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEIDIETIG 10
 |||||:|:|
 Db 158 geidmetvny 166

RESULT 10

AAY32318
 ID AAY32318 standard; Protein: 223 AA.

XX AAY32318;

DT 28-FEB-2000 (first entry)

DE Corn beta-carotene hydroxylase.

XX Beta-carotene hydroxylase; corn; maize; transgenic plant;

KM carotenoid; zeaxanthin; pigment.

XX Zea mays;

XX WO955887-A2.

XX WO955887-A2.

XX 04-NOV-1999.

XX 16-APR-1999; 99WO-US08384.

XX 24-APR-1998; 98US-0083042.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Caboon RE, Kinney AJ, Pearlstein RW, Williams ME;

XX WPI; 2000-062037/05.

XX N-PSDB; AA234967.

XX Novel carotenoid biosynthesis enzyme polynucleotides and polypeptides

used to identify inhibitors and to modulate expression of the enzyme

Clalm 10; Page 40-41; 61pp; English.

This is the deduced amino acid sequence of the C-terminal three quarters of a corn beta-carotene hydroxylase derived from isolated cDNA clones (see AA334967). The enzyme converts beta-carotene into zeaxanthin. The invention provides novel carotenoid biosynthesis enzymes, specifically beta-carotene hydroxylase, lycopene cyclase and lycopene epsilon cyclase (see AA332318-26 and AA332386-88), and polynucleotides encoding them (see AA334967-74). The enzymes may be prepared recombinantly and used to raise antibodies, used for detecting the enzymes *in situ* or *in vitro*. The polynucleotides may be used to create transgenic plants in which the enzymes are present at higher or lower levels than normal, or in cell types or developmental processes where they are not normally found. This would alter the level of alpha-carotene or beta-carotene in those cells. Zeaxanthin gives a consumer-preferred darker colour to yolks and poultry products, and so is a desirable poultry feed additive. Blocking of beta-carotene hydroxylase activity may create a high beta-carotene corn which may be valuable for human consumption. The enzymes can also be used as targets to facilitate the design and/or identification of inhibitors of those enzymes that may be useful as herbicides. This is desirable because inhibition of any of the enzymes could lead to an inhibition of plant growth.

Query Match	59.4%	Score 38;	DB 21;	Length 223;
Best Local Similarity	60.0%;	Pred. No. 16;		
Matches 6; Conservative		Mismatches 1;	Indels 0;	Gaps 0;

```
QY      2 GEIDIETIG 11
        ||: :|||:|
Db      24 gevpietlg 33
```

RESULT	11
AA644125	
ID	AA644125 standard; Protein; 275 AA.

AC AAG44125;

DF 18-OCT-2000 (first entry)

DE: Arabidopsis thaliana protein fragment SEQ ID NO: 55231

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000

PF 25-FEB-2000; 2000EP-0301439.

PR	25-FEB-1999;	9905-0121825
PR	05-MAR-1999;	9905-0123180
PR	09-MAR-1999;	9905-0123548
PR	23-MAR-1999;	9905-0125788
PR	25-MAR-1999;	9905-0126264
PR	29-MAR-1999;	9905-0126785
PR	01-APR-1999;	9905-0127462
PR	06-APR-1999;	9905-0128234
PR	08-APR-1999;	9905-0128714
PR	16-APR-1999;	9905-0129645
PR	19-APR-1999;	9905-0130077
PR	21-APR-1999;	9905-0130449
PR		9905-0130510

PR	23-APR-1999;	9905-0130691;
PR	28-APR-1999;	9905-0130448;
PR	30-APR-1999;	9905-0133049;
PR	30-APR-1999;	9905-0133407;
PR	04-MAY-1999;	9905-0133404;
PR	05-MAY-1999;	9905-0133485;
PR	06-MAY-1999;	9905-0132486;
PR	06-MAY-1999;	9905-0133487;
PR	07-MAY-1999;	9905-0133863;
PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134218;
PR	14-MAY-1999;	9905-0134322;
PR	14-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134368;
PR	19-MAY-1999;	9905-0134541;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135353;
PR	24-MAY-1999;	9905-0133629;
PR	25-MAY-1999;	9905-0136021;
PR	27-MAY-1999;	9905-0136392;
PR	28-MAY-1999;	9905-0136782;
PR	01-JUN-1999;	9905-0137222;
PR	03-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137502;
PR	07-JUN-1999;	9905-0137724;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138360;
PR	10-JUN-1999;	9905-0138547;
PR	14-JUN-1999;	9905-0138311;
PR	16-JUN-1999;	9905-0133453;
PR	16-JUN-1999;	9905-0133452;
PR	18-JUN-1999;	9905-0133458;
PR	18-JUN-1999;	9905-0133476;
PR	18-JUN-1999;	9905-0133454;
PR	18-JUN-1999;	9905-0133455;
PR	18-JUN-1999;	9905-0133456;
PR	18-JUN-1999;	9905-0133457;
PR	18-JUN-1999;	9905-0133458;
PR	18-JUN-1999;	9905-0133750;
PR	18-JUN-1999;	9905-0133763;
PR	21-JUN-1999;	9905-0133817;
PR	22-JUN-1999;	9905-0140359;
PR	23-JUN-1999;	9905-0140353;
PR	23-JUN-1999;	9905-0140354;
PR	24-JUN-1999;	9905-0140695;
PR	28-JUN-1999;	9905-0140823;
PR	30-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142151;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;

PR	01-JUL-1999;	9905-0144884;
PR	21-JUL-1999;	9905-0144818;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147305;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150564;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151086;
PR	27-AUG-1999;	9905-0151087;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151338;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0153363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	24-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155569;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157535;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159294;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159537;

Query Match	Best Local	Similarity	57.8%	Score 37;	DB 21;	Length 275;
Matches	5;	Conservative	4;	Mismatches	1;	Indels
						Gaps
						0;
Qy	3	EVIDENCE	12			
	1:1:1:1:1					
Db	75	evdqvqivgn	84			
RESULT	12					
AAAG44124						
ID	AAAG44124	standard; Protein; 324 AA.				
XX	AAAG44124;					
XX						
XX	18-OCT-2000	(first entry)				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55230.					
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.					
XX	Arabidopsis thaliana.					
OS	EP1033405-A2.					
PN						
XX	06-SEP-2000.					
PD						
XX	25-FEB-2000; 2000EP-0301439.					
XX						
PR	25-FEB-1999;	99US-0121825.				
PR	05-MAR-1999;	99US-0123180.				
PR	09-MAR-1999;	99US-0123548.				
PR	23-MAR-1999;	99US-0125788.				
PR	25-MAR-1999;	99US-0126264.				
PR	29-MAR-1999;	99US-0126785.				
PR	01-APR-1999;	99US-0127462.				
PR	06-APR-1999;	99US-0128234.				
PR	08-APR-1999;	99US-0128714.				
PR	16-APR-1999;	99US-0129845.				
PR	21-APR-1999;	99US-0130077.				
PR	23-APR-1999;	99US-0130444.				
PR	23-APR-1999;	99US-0130510.				
PR	28-APR-1999;	99US-0130891.				
PR	30-APR-1999;	99US-0131449.				
PR	30-APR-1999;	99US-0132048.				
PR	04-MAY-1999;	99US-0132407.				
PR	04-MAY-1999;	99US-0132484.				
PR	05-MAY-1999;	99US-0132485.				
PR	06-MAY-1999;	99US-0132486.				

PR 06-MAY-1999;	990US-0132487.
PR 07-MAY-1999;	990US-0132863.
PR 11-MAY-1999;	990US-0134256.
PR 14-MAY-1999;	990US-0134218.
PR 14-MAY-1999;	990US-0134219.
PR 14-MAY-1999;	990US-01343271.
PR 14-MAY-1999;	990US-0134370.
PR 18-MAY-1999;	990US-01343768.
PR 19-MAY-1999;	990US-0134941.
PR 20-MAY-1999;	990US-0135124.
PR 21-MAY-1999;	990US-0135353.
PR 24-MAY-1999;	990US-0135629.
PR 25-MAY-1999;	990US-0136021.
PR 27-MAY-1999;	990US-0136392.
PR 28-MAY-1999;	990US-0136782.
PR 01-JUN-1999;	990US-0137222.
PR 03-JUN-1999;	990US-0137528.
PR 04-JUN-1999;	990US-0137502.
PR 07-JUN-1999;	990US-0137724.
PR 08-JUN-1999;	990US-0138094.
PR 10-JUN-1999;	990US-0138540.
PR 10-JUN-1999;	990US-0138847.
PR 14-JUN-1999;	990US-0139119.
PR 16-JUN-1999;	990US-0139452.
PR 16-JUN-1999;	990US-0139453.
PR 17-JUN-1999;	990US-0139492.
PR 18-JUN-1999;	990US-0139454.
PR 18-JUN-1999;	990US-0139455.
PR 18-JUN-1999;	990US-0139456.
PR 18-JUN-1999;	990US-0139457.
PR 18-JUN-1999;	990US-0139458.
PR 18-JUN-1999;	990US-0139459.
PR 18-JUN-1999;	990US-0139460.
PR 18-JUN-1999;	990US-0139461.
PR 18-JUN-1999;	990US-0139462.
PR 18-JUN-1999;	990US-0139463.
PR 18-JUN-1999;	990US-0139750.
PR 21-JUN-1999;	990US-0139817.
PR 22-JUN-1999;	990US-0139899.
PR 23-JUN-1999;	990US-0140353.
PR 23-JUN-1999;	990US-0140354.
PR 24-JUN-1999;	990US-0140695.
PR 28-JUN-1999;	990US-0140823.
PR 29-JUN-1999;	990US-0140991.
PR 30-JUN-1999;	990US-0141287.
PR 01-JUL-1999;	990US-0141842.
PR 01-JUL-1999;	990US-0142154.
PR 02-JUL-1999;	990US-0142055.
PR 06-JUL-1999;	990US-0142390.
PR 08-JUL-1999;	990US-0142803.
PR 09-JUL-1999;	990US-0142920.
PR 12-JUL-1999;	990US-0142977.
PR 13-JUL-1999;	990US-0143542.
PR 14-JUL-1999;	990US-0143624.
PR 15-JUL-1999;	990US-0144005.
PR 16-JUL-1999;	990US-0144085.
PR 16-JUL-1999;	990US-0144086.
PR 19-JUL-1999;	990US-0144325.
PR 19-JUL-1999;	990US-0144331.
PR 19-JUL-1999;	990US-0144332.
PR 19-JUL-1999;	990US-0144333.
PR 19-JUL-1999;	990US-0144334.
PR 19-JUL-1999;	990US-0144335.
PR 20-JUL-1999;	990US-0144352.
PR 20-JUL-1999;	990US-0144352.
PR 20-JUL-1999;	990US-0144884.
PR 20-JUL-1999;	990US-0144884.
PR 21-JUL-1999;	990US-0145086.
PR 21-JUL-1999;	990US-0145086.
PR 22-JUL-1999;	990US-0145085.
PR 22-JUL-1999;	990US-0145087.
PR 22-JUL-1999;	990US-0145089.
PR 22-JUL-1999;	990US-0145089.
PR 22-JUL-1999;	990US-0145192.
PR 23-JUL-1999;	990US-0145145.
PR 23-JUL-1999;	990US-0145218.
PR 23-JUL-1999;	990US-0145224.
PR 26-JUL-1999;	990US-0145276.
PR 27-JUL-1999;	990US-0145813.
PR 27-JUL-1999;	990US-0145918.
PR 27-JUL-1999;	990US-0145919.
PR 28-JUL-1999;	990US-0145951.
PR 02-AUG-1999;	990US-0146386.
PR 02-AUG-1999;	990US-0146388.
PR 02-AUG-1999;	990US-0146389.
PR 03-AUG-1999;	990US-0147038.
PR 04-AUG-1999;	990US-0147204.
PR 04-AUG-1999;	990US-0147302.
PR 05-AUG-1999;	990US-0147192.
PR 05-AUG-1999;	990US-0147260.
PR 06-AUG-1999;	990US-01474303.
PR 06-AUG-1999;	990US-0147436.
PR 09-AUG-1999;	990US-0147493.
PR 09-AUG-1999;	990US-0147935.
PR 10-AUG-1999;	990US-0147935.
PR 10-AUG-1999;	990US-0148171.
PR 11-AUG-1999;	990US-0148319.
PR 12-AUG-1999;	990US-0148341.
PR 13-AUG-1999;	990US-0148565.
PR 13-AUG-1999;	990US-0148684.
PR 16-AUG-1999;	990US-0149368.
PR 17-AUG-1999;	990US-0149175.
PR 18-AUG-1999;	990US-0149426.
PR 20-AUG-1999;	990US-0149722.
PR 20-AUG-1999;	990US-0149723.
PR 20-AUG-1999;	990US-0149929.
PR 23-AUG-1999;	990US-0149902.
PR 23-AUG-1999;	990US-0149930.
PR 25-AUG-1999;	990US-0150566.
PR 26-AUG-1999;	990US-0150884.
PR 27-AUG-1999;	9

PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 324;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
3 EIDITION 12
124 evdvlqvlgn 133

RESULT 13
AAG44123
ID AAG44123 standard; Protein; 336 AA.
XX AAG44123;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55229.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127462.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129645.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130044.
PR 23-APR-1999: 99US-0130510.
PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.

PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137502.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139753.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142380.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.

```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```

```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match : 57.88; Score 37; DB 21; Length 336;
Best Local Similarity 50.08; Pred. No. 41;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 EIDIFITGN 12
Db 136 evdvlqvign 145

RESULT 14
AAE00115
ID AAE00115 standard; Protein; 222 AA.
AC AAE00115;
AC AAE00115;
DT 31-MAY-2001 (first entry)
XX
DE Bacillus licheniformis RP-II protease variant (T109R).
XX
KM Residual protease II; RP-II; additive; cleaning composition; detergent;
XX mutant; muten; variant.
OS Bacillus licheniformis.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Msc-difference 109
FT /note= "Wild type Thr substituted with Arg"
XX
PN W0200116285-A2.
XX
PD 08-MAR-2001.
XX
XX
PF 31-AUG-2000; 2000WO-DK00476.
XX
PR 31-AUG-1999; 99DK-0001212.
PR 20-OCT-1999; 99DK-0001500.
XX
XX (NOVO ) NOVOZYMES AS.
XX
XX Norregaard-madsen M, Rahbek Ostergaard P, Voge Christensen CB;
PI Plensted Lassen S;
XX
XX WPI; 2001-226680/23.
XX
XX Novel RP-II type protease and its variants useful as constituents in
XX detergent compositions, additives and cleaning compositions
XX
XX Example 3; Page -; 132pp; English.
XX
XX The present sequence is Bacillus licheniformis (BLC) RP-II protease
XX variant (T109R).
XX The patent discloses RP-II (Residual protease II) protein, their
XX corresponding nucleic acid sequences and variants. RP-II protease
XX is useful as a constituent in additives, detergent compositions and
XX other cleaning compositions, optionally in combination with other
XX enzymes such as proteases, lipases, cellulases, amylases, peroxidases
XX or oxidases. The variants of RP-II have improved properties such as
XX substrate specificities, catalytic rate, stability, especially towards
XX the action of proteolytic enzymes and improved resistance towards
XX peroxidase.
XX Note: This sequence is not shown in the specification but is derived
XX from the mature B. licheniformis RPII protein [SEQ ID NO: 2] shown in
XX page 101-102 of the specification (AAE00011). The specification also
XX refers to following variants: (a) V1F, (b) D7G+T125+E152G+M182I,

```


CC (c) S119A+S144T+Q157R+S175I, (d) H141A, (e) E152A+Q122R, (f) N182T+C185A,
 CC (g) E152A+V144L and (h) E104K+Q204R. However the positions of Val in (a)
 CC and (g), Asn residue in (b) and (f), Ser residues in (c), His residue in
 CC (d), Gln residue in (e) and (h) do not match with the B. licheniformis
 CC native wild type RPII protease sequence.

SO Sequence 222 AA;

Query Match 56.2%; Score 36; DB 22; Length 222;
 Best Local Similarity 58.3%; Pred. No. 39;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GEIDIIETIGNR 13
 | | | | |
 Db 98 galelsep19r 109

RESULT 15

AAB70788
 AAB70788 standard; Protein; 419 AA.

AAB70788;

25-MAY-2001 (first entry)

N. magadali bacteriophage phi-CH1 6-methyltransferase.

Genome; halophilic; polyhydroxybutyrate; inducible expression;
 6-methyltransferase.

Bacteriophage phi-CH1.

Natrialba magadii.

DE19937719-A1.

22-FEB-2001.

10-AUG-1999; 99DE-1037719.

10-AUG-1999; 99DE-1037719.

(LDB1/) LDB1TZ W.

Witte A, Baranyi U, Klein R;

WPI; 2001-245930/26.

N-PSDB; AAF61284.

New nucleic acid from phage phiCH1, used to create vectors for
 expressing proteins and polymers in halophilic Archaea -

Claim 9; Page 41-42; 72pp; German.

This invention describes a novel isolated nucleic acid (I) from the
 genome of phage phi-CH1 of Natrialba magadii. The genome of phi-CH1 is a
 combination of 48300 and 10198 base pair sequences (S1 and S2), both
 fully defined in the specification. The invention also describes a
 novel (1) recombinant vector (RV1) containing at least one copy of (I);
 (2) recombinant vector (RV2) containing the ori of phi-CH1; (3) cell
 transformed with RV1 or RV2; (4) isolated polypeptide (II) encoded by
 (I); N. magadii cells free from the prophage of phi-CH1; (5) phage
 variants having lytic properties different from those of wild-type
 phi-CH1 and having genomic sequences at least 70 % homologous with the
 phi-CH1 genome; (6) use of phi-CH1 as gene transfer vector; and (7) use
 of halophilic Archaea for production of proteins and other polymers.
 Vectors that contain (I) are used to transform halophilic Archaea,
 specifically N. magadii, for production of proteins and other polymers
 (e.g. poly(hydroxybutyrate)). Vectors containing (I) allow controlled,
 inducible expression of compounds in Archaea.

Sequence 419 AA;

Query Match 56.2%; Score 36; DB 22; Length 419;
 Best Local Similarity 54.5%; Pred. No. 81;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 3 EIDIIETIGNR 13
 | | | | |
 Db 202 did1set19r 212

Search completed: October 11, 2001, 15:51:38
 Job time: 201 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:48:37 ; Search time 19.14 seconds
(without alignments)
13.985 Million cell updates/sec

Title: US-09-596-101A-1
Perfect score: 64
Sequence: 1 SGEIDIIETIGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 197339 segs, 20590346 residues
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCPMUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	78.1	306	2	US-08-824-707-2
2	39	60.9	226	2	US-08-712-072C-4
3	39	60.9	226	2	US-08-737-526-4
4	39	60.9	292	4	US-09-098-580-4
5	37	57.8	262	1	US-08-392-828C-37
6	37	57.8	262	3	US-09-330-945-37
7	37	57.8	321	2	US-08-712-072C-3
8	35	54.7	430	1	US-08-476-008-44
9	35	54.7	430	1	US-08-306-063-44
10	35	54.7	430	1	US-08-833-485-44
11	35	54.7	430	1	US-09-137-440-44
12	34	53.1	654	1	US-08-392-828C-2
13	34	53.1	654	3	US-09-330-945-2
14	34	53.1	1119	4	US-09-396-651B-2
15	34	53.1	2409	6	5180808-2
16	33	51.6	394	1	US-07-621-193A-5
17	33	51.6	394	1	US-08-018-489C-5
18	33	51.6	434	2	US-09-008-962-1
19	33	51.6	434	2	US-08-675-507-1
20	33	51.6	434	3	US-09-213-205-1
21	33	51.6	761	1	US-07-906-395-2
22	33	51.6	761	1	US-08-192-632-2
23	33	51.6	761	1	US-08-710-676-2
24	33	51.6	761	1	US-09-099-902B-2
25	33	51.6	761	5	PCT-US93-06080-2
26	33	51.6	887	1	US-07-867-106-3
27	33	51.6	4551	3	US-09-330-878-1

28	32	50.0	77	2	US-08-343-443B-13	Sequence 13, Appl
29	32	50.0	114	1	US-08-031-399-3	Sequence 3, Appl
30	32	50.0	114	1	US-08-393-305-6	Sequence 6, Appl
31	32	50.0	114	1	US-08-726-817-6	Sequence 6, Appl
32	32	50.0	114	1	US-08-504-042-3	Sequence 3, Appl
33	32	50.0	114	2	US-08-725-969-6	Sequence 6, Appl
34	32	50.0	114	2	US-08-794-524-6	Sequence 6, Appl
35	32	50.0	114	4	US-09-189-193-6	Sequence 6, Appl
36	32	50.0	114	5	PCT-US94-03793-3	Sequence 3, Appl
37	32	50.0	122	1	US-08-300-903A-3	Sequence 3, Appl
38	32	50.0	162	1	US-08-031-399-2	Sequence 2, Appl
39	32	50.0	162	1	US-08-393-305-5	Sequence 5, Appl
40	32	50.0	162	1	US-08-284-393B-9	Sequence 9, Appl
41	32	50.0	162	1	US-08-726-817-5	Sequence 5, Appl
42	32	50.0	162	1	US-08-504-042-2	Sequence 2, Appl
43	32	50.0	162	2	US-08-725-969-5	Sequence 5, Appl
44	32	50.0	162	2	US-08-794-524-5	Sequence 5, Appl
45	32	50.0	162	4	US-09-189-193-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-824-707-2
Sequence 2, Application US/08824707
Patent No. 5919688
GENERAL INFORMATION:
APPLICANT: Ferrer, Pau
APPLICANT: Diets, Ivan
APPLICANT: Hedegaard, Lisbeth
APPLICANT: Halkier, Torben
APPLICANT: Asenjo, Juan
APPLICANT: Savva, Demetris
TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5919688o No. 5919688dsk of No. 5919688h America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,707
FILING DATE: 14-April-1997
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4290,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-824-707-2

Query Match: 78.1%; Score 50; DB 2; Length 306;
Best Local Similarity: 75.0%; Pred. No. 0.09;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SGEIDIIETIGN 12

APPLICATION NUMBER: 08/737,526
FILING DATE: 08-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Valeta, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-098-580-4

Query Match 60.8%; Score 39; DB 4; Length 292;
Best Local Similarity 77.8%; Pred. NO. 9.3;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GSEIDIIETI 10
|||||:|:|
DB 158 GSEIDIMETV 166

RESULT 5
US-08-392-828C-37
Sequence 37, Application US/08392828C
Patent No. 5795962
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAKI
APPLICANT: MUTA, TATSUSHI
APPLICANT: SEKI, NORIKI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,828C
FILING DATE: 28-FEB-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein

LOCATION: 1..262
OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)"
US-08-392-828C-37

Query Match 57.8%; Score 37; DB 1; Length 262;
Best Local Similarity 63.6%; Pred. NO. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GSEIDIIETIG 11
|||||:|:|
DB 130 GSEIDVMEARG 140

RESULT 6
US-09-330-945-37
Sequence 37, Application US/09330945
Patent No. 6077946
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAKI
APPLICANT: MUTA, TATSUSHI
APPLICANT: SEKI, NORIKI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIBEAULT, LLP
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: FJN-032DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..262
OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)"
US-09-330-945-37

Query Match 57.8%; Score 37; DB 3; Length 262;
Best Local Similarity 63.6%; Pred. NO. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GSEIDIIETIG 11
|||||:|:|

Db 130 SGEIDVMEARG 140

RESULT 7

US-08-712-072C-3

Sequence 3, Application US/08712072C
Patent No. 5925541

GENERAL INFORMATION:

APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng

TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amster, Rothstein & Eberstein

STREET: 90 Park Avenue

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/712,072C

FILING DATE: 11-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Elizabeth A.

REGISTRATION NUMBER: 39,911

REFERENCE/DOCKET NUMBER: 63475/97

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-5995

TELEFAX: (212) 286-0854 or 286-0082

TELEX: TWX 710-581-4766

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: el3b, Bacillus circulans

08-712-072C-3

QY 1 SGEIDIETIG 11

Db 189 SGEIDVMEARG 199

RESULT 8

US-08-476-008-44

Sequence 44, Application US/08476008

Patent No. 5627061

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Padgett, Stephen R.

APPLICANT: Stallings, William C.

TITLE OF INVENTION: 5-Ethylpyruvylshikimate-3-Phosphate Synthases

TITLE OF INVENTION: 5-Ethylpyruvylshikimate-3-Phosphate Synthases

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,008

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,063

FILING DATE: 13-SEP-1994

APPLICATION NUMBER: US 07/749,611

FILING DATE: 28-AUG-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,537

FILING DATE: 31-AUG-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10660)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6047

TELEFAX: (314)537-6099

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 430 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-476-008-44

QY 1 SGEIDIETIG 11

Db 265 SGEIDVMEARG 275

RESULT 9

US-08-306-063-44

Sequence 44, Application US/08306063

Patent No. 5633435

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Padgett, Stephen R.

APPLICANT: Stallings, William C.

TITLE OF INVENTION: 5-Ethylpyruvylshikimate-3-Phosphate Synthases

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-44

Query Match 54.7%; Score 35; DB 1; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11

DB 265 SGIDIVKMG 275

RESULT 10
US-08-833-485-44
Sequence 44, Application US/08833485
Patent No. 5804425
GENERAL INFORMATION:
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-44

Query Match 54.7%; Score 35; DB 1; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11

DB 265 SGIDIVKMG 275

RESULT 11
US-09-137-440-44
Sequence 44, Application US/09137440
Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-44

Query Match 54.7%; Score 35; DB 4; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDITE 11
11111111
265 SGEIDITE 275

RESULT 12
US-08-392-828C-2
Sequence 2, Application US/08392828C
Patent No. 5785962
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAOKI
APPLICANT: MOTA, TATSUSHI
APPLICANT: SEKI, NORIHI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,828C
FILING DATE: 28-FEB-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-392-828C-2

Query Match 53.1%; Score 34; DB 1; Length 654;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITE 8

Db 116 SGEIDITE 123
11111111

RESULT 13
US-09-330-945-2
Sequence 2, Application US/09330945
Patent No. 6077946
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAOKI
APPLICANT: MOTA, TATSUSHI
APPLICANT: SEKI, NORIHI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: FJN-032DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-330-945-2

Query Match 53.1%; Score 34; DB 3; Length 654;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITE 8
11111111
Db 116 SGEIDITE 123

RESULT 14
US-09-396-651B-2
Sequence 2, Application US/09396651B
Patent No. 6225076
GENERAL INFORMATION:
APPLICANT: Darst, Seth A
APPLICANT: Zhang, Gongyi
APPLICANT: Campbell, Elizabeth
APPLICANT: Minakim, Leonid
APPLICANT: Severinov, Konstantin
TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
TITLE OF INVENTION: OF USE THEREOF


```

1 FILE REFERENCE: 600-1-258
2 CURRENT APPLICATION NUMBER: US/09/396,651B
3 CURRENT FILING DATE: 1999-09-15
4 NUMBER OF SEQ ID NOS: 4
5 SOFTWARE: PatentIn Ver. 2.0
6 SEQ ID NO 2
7
8 LENGTH: 1119
9
10 TYPE: PRT
11
12 ORGANISM: Thermus aquaticus
13
14 FEATURE:
15
16 NAME/KEY: SITE
17 LOCATION: (695)..(696)
18
19 OTHER INFORMATION: Any amino acids can be at these two positions.
20
21 US-09-396-651B-2

```

Query Match	53.1%	Score 34:	DB 4:	length 1119;
Best Local Similarity	54.5%	Pred. No.	3.6e+02;	
--Matches	6;	Conservative	3;	Mismatches 2; Indels 0; Gaps 0;

```

      3 EIDIETIGNR 13
      1:| |: |||
Db    321 EVDDIDHLGNR 331

```

RESULT 15
5180808-2
Patent No. 5180808
Applicant: RUOSLAHTI, EERIK I.
Title of Invention: VERSICAN CORE PROTEIN, NUCLEIC ACID
SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN
ANTIBODIES, AND METHODS OF DETECTING THE SAME
Number of Sequences: 4
Current Application Number: 1
Application Number: US/07/441,179
Filing Date: 27-NOV-1989
Seq ID No.: 2
Length: 2409
5180808-2

Query Match	53.1%	Score 34:	DB 6:	Length 2409:
Best Local Similarity	55.6%	Pred. No.	8.7e+02:	
Matches	5:	Conservative	4:	Mismatches 0: Indels 0: Gaps 0:

QY 1 SGEIDIET 9
111:11::
1446 SGEVDIVDS 1454

Search completed: October 11, 2001, 15:52:04
Job time: 207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:52:08 ; Search time 15.16 Seconds
(without alignments)
29.375 Million cell updates/sec

Title: US-09-596-101A-1

Perfect score: 64
Sequence: 1 SGEIDITITGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 93435 segs, 34255486 residues

Cal number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	62.5	782	1 PRTP_HSVT2	O9wrl6 herpesvirus
2	40	62.5	806	1 MK07_MOUSE	O9wz88 mus musculus
3	40	62.5	815	1 MK07_HUMAN	Q13164 homo sapien
4	40	62.5	1302	1 RPOB_SPTCI	P47767 spitioplama
5	40	62.5	1342	1 RPOB_BUCAI	P57166 buchiera ap
6	40	62.5	1342	1 RPOB_BUCAP	P41184 buchiera ap
7	40	62.5	1342	1 RPOB_ECOLI	P00575 escherichia
8	40	62.5	1342	1 RPOB_SALTY	P06173 salmonella
9	40	62.5	1343	1 RPOB_HAEIN	P43738 haemophilus
10	40	62.5	1392	1 RPOB_NEIMA	P57009 neisseria m
11	40	62.5	1392	1 RPOB_NEIMB	O59622 neisseria m
12	40	62.5	1655	1 N188_YEAST	P52593 saccharomyc
13	39	60.9	286	1 GDB_RHOMR	P45798 rhodothermu
14	39	60.9	540	1 TCPO_CANAL	P47828 candida alb
15	39	60.9	665	1 CNG_DROME	Q24278 drosophila
16	38	59.4	371	1 MALX_ECOLI	P02919 escherichia
17	38	59.4	1103	1 RPOB_CYPAP	P51252 cyanophora
18	38	59.4	1143	1 RPOB_PORPU	P51252 porphyra pu
19	37	57.8	309	1 AGAR_STRCO	P07883 streptomyce
20	37	57.8	682	1 E13B_BACCI	P23903 bacillus ci
21	37	57.8	1375	1 RPOB_COXBU	O87903 coxiella bu
22	37	57.8	1492	1 AT7A_RAT	P70705 rattus norv
23	36	56.2	446	1 YMA2_YEAST	O00359 cochlidiolu
24	36	56.2	737	1 YMA2_YEAST	O04263 saccharomyc
25	36	56.2	1096	1 RPOB_GUTTH	O78485 guillardia
26	36	56.2	1194	1 APAP_HUMAN	O14727 homo sapien
27	35	55.5	546	1 TCPO_SCHPO	P78921 schizosach
28	35	54.7	151	1 VG14_BPP22	O01075 bacterioph
29	35	54.7	261	1 YY00_MYCLE	O49741 mycobacteri
30	35	54.7	313	1 K1PE_HAEIN	P44330 haemophilus
31	35	54.7	314	1 MTRH_METMA	P80650 methanosarc
32	35	54.7	393	1 DOM3_CABEL	Q10660 caenorhabdi
33	35	54.7	430	1 AROA_STRAU	O05615 staphylococ

34	35	54.7	546	1 TRCD_ARCFU	O28668 archaeoglob
35	35	54.7	690	1 CYG3_HUMAN	Q02108 homo sapien
36	35	54.7	690	1 CYG3_RAT	P19686 rattus norv
37	35	54.7	691	1 CYG3_BOVIN	P19687 bos taurus
38	35	54.7	728	1 ARPA_ECOLI	P23325 escherichia
39	35	54.7	788	1 SYFB_THERMA	O9wz89 thermotoga
40	35	54.7	897	1 DYHC_HUMAN	Q14204 homo sapien
41	35	54.7	2511	1 FAS_CHICK	P12276 gallus gall
42	35	54.7	4344	1 DYHC_EMENT	P45444 emericella
43	35	54.7	4349	1 DYHC_FUSSO	P78716 fusarium so
44	35	54.7	4367	1 DYHC_NEUCR	P45443 neurospora
45	35	54.7	4639	1 DYHC_DROME	P37276 drosophila

ALIGNMENTS

```

RESULT 1
PRTP_HSVT2
ID PRTP_HSVT2 STANDARD; PRT; 782 AA.
AC O9wrl6;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE PROCESSING AND TRANSPORT PROTEIN.
GN PRTP.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RX MEDLINE=99319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
RT processing and transport protein, and the major DNA binding protein."
RL Virus Res. 60:123-136(1999).
CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION
(BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF084543; AAD42934.1; -
DR InterPro: IPR000501; -
DR Pfam: PF01366; PRTP. 1.
KW Capsid assembly.
SQ
SEQUENCE 782 AA; 87392 MW; 4DE40314E5BABB9B CRC64;

Query Match
Best Local Similarity 62.5%; Score 40; DB 1; Length 782;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDITITGNR 13
DB 236 EDDIVETVGR 246

RESULT 2
MK07_MOUSE
ID MK07_MOUSE STANDARD; PRT; 806 AA.
AC O9wz88;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)

```

01-OCT-2000 (Rel. 40, last annotation update)
 DE MITOGEN-ACTIVATED PROTEIN KINASE 7 (EC 2.7.1.1-) (EXTRACELLULAR SIGNAL-
 REGULATED KINASE 5) (ERK-5) (BMK1 KINASE).
 GN MAPK7 OR ERK5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamakura S., Moriyuchi T., Nishida E.;
 RT Activation of the protein kinase ERK5/BMK1 by receptor tyrosine
 RT kinases: identification and characterization of a signaling pathway to
 the nucleus.";
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND
 CC NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS (BY SIMILARITY).
 CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN
 CC TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.
 CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES, WHEN
 CC THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY
 CC ROLE, IS ABSENT (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB019373; BAA82039.1; -;
 CC DR HSSP: P27703; ZERK.
 CC DR MGD: MGI:1346347; Mapk7.
 CC DR InterPro: IPR000719; -;
 CC DR InterPro: IPR002290; -;
 CC DR PROSITE: PS01351; MAPK: 1.
 CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC KM Transphosphatase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 CC Phosphorylation.
 CC FT DOMAIN 55 347 PROTEIN KINASE.
 CC FT DOMAIN 434 465 PRO-RICH 1.
 CC FT DOMAIN 521 524 POLY-ARG.
 CC FT NP_BIND 61 69 ATP (BY SIMILARITY).
 CC FT BINDING 84 84 ATP (BY SIMILARITY).
 CC FT ACT_SITE 182 182 BY SIMILARITY.
 CC FT MOD_RES 219 219 PHOSPHORYLATION (ACTIVATES THE KINASE)
 CC FT MOD_RES 221 221 (BY SIMILARITY).
 CC FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE)
 CC FT MOD_RES 221 221 (BY SIMILARITY).
 CC SQ SEQUENCE 806 AA; 87732 MW; E7CC41CABDDE0633 CRC64;
 FT
 Query Match 62.5%; Score 40; DB 1; Length 806;
 Best local Similarity 80.0%; Pred. No. 18;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

01-NOV-1997 (Rel. 35, last annotation update)
 DE MITOGEN-ACTIVATED PROTEIN KINASE 7 (EC 2.7.1.1-) (EXTRACELLULAR SIGNAL-
 REGULATED KINASE 5) (ERK-5) (BMK1 KINASE).
 GN MAPK7 OR ERK5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhou G., Bao Z.Q., Dixon J.E.;
 RT "Components of a new human protein kinase signal transduction
 RT pathway.";
 RT J. Biol. Chem. 270:12665-12669(1995).
 CC -1- FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND
 CC NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS.
 CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY ADULT TISSUES. ABUNDANT IN
 CC HEART, PLACENTA, LUNG, KIDNEY AND SKELETAL MUSCLE. NOT DETECTABLE
 CC IN LIVER.
 CC -1- DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN
 CC TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.
 CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES, WHEN
 CC THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY
 CC ROLE, IS ABSENT.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U25278; AAA81381.1; -;
 CC DR EMBL: U29725; AAA82931.1; -;
 CC DR EMBL: U29726; AAA82932.1; -;
 CC DR EMBL: U29727; AAA82933.1; -;
 CC DR HSSP: P24941; IAO1.
 CC DR MIM: 602521; -;
 CC DR InterPro: IPR000719; -;
 CC DR InterPro: IPR002290; -;
 CC DR Pfam: PF00069; pkinase; 1.
 CC DR PROSITE: PS01351; MAPK: 1.
 CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC KM Transphosphatase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
 CC Phosphorylation.
 CC FT DOMAIN 54 346 PROTEIN KINASE.
 CC FT DOMAIN 337 340 POLY-ALA.
 CC FT DOMAIN 433 464 PRO-RICH 1.
 CC FT DOMAIN 520 523 POLY-ARG.
 CC FT NP_BIND 577 699 ATP (BY SIMILARITY).
 CC FT BINDING 60 68 ATP (BY SIMILARITY).
 CC FT ACT_SITE 83 83 BY SIMILARITY.
 CC FT ACT_SITE 181 181 PHOSPHORYLATION (ACTIVATES THE KINASE)
 CC FT MOD_RES 218 218 (BY SIMILARITY).
 FT

FT MOD.RES 220 220 PHOSPHORYLATION (ACTIVATES THE KINASE)
 FT CONFLICT 19 32 (BY SIMILARITY).
 FT CONFLICT 609 609 AREGTRPHRCIS -> GPVKEPATASVA (IN
 FT SEQUENCE 815 AA: 88636 MW: 379AD69803207CCF CRC64;
 REF. 2).
 L -> V (IN REF. 2).

Query Match 62.5%; Score 40; DB 1; Length 815;
 Best Local Similarity 80.0%; Pred. No. 18;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDIIETIGN 12
 I:|||||
 DB 53 EYEIIETIGN 62

RESULT 4
 PROB.SPICI STANDARD; PRT; 1302 AA.

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).

OS Spiroplasma citri.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Spiroplasmataceae; Spiroplasma.
 OX NCBI_TaxID=2133;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ROZHP;
 RX MEDLINE=96257200; PubMed=8675039;

RT Laigret F., Gaurivaud P., Bove J.;

RT "The unique organization of the rpoB region of Spiroplasma citri: a
 restriction and modification system gene is adjacent to rpoB.";

RL Gene 171:95-98(1996).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 RNA(N).

CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U25815; AAC44217.1; -

DR InterPro: IPR001572; -

DR Pfam: PF00562; RNA_POL_B; 1.

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KM Transference: Transcription: DNA-directed RNA polymerase.

SO SEQUENCE 1302 AA; 146533 MW; 8D971C23EB9FEB2F CRC64;

OY 2 GEIDIIETIGNR 13
 I:|:|:|
 DB 475 GEVDIDIDLGNR 486

Query Match 62.5%; Score 40; DB 1; Length 1302;
 Best Local Similarity 58.3%; Pred. No. 30;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
 PROB.BUCAI STANDARD; PRT; 1342 AA.

AC P57146;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
 GN PROB OR B0034.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-TOKYO 1998;

RX MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. Aps.";

RL Nature 407:81-86(2000).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 RNA(N).

CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AP001118; BAB12761.1; -

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KM Transference: Transcription: DNA-directed RNA polymerase.

SO SEQUENCE 1342 AA; 151467 MW; 95252459873DF940 CRC64;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|

DB 440 GEVDIDIDLGNR 451

CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -1- CAUTION: REF.6 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00339; CAA23625.1; -
DR EMBL: U00006; AAC43085.1; -
DR EMBL: AE000472; AAC76961.1; -
DR EMBL: U76222; AAB18647.1; -
DR EMBL: V00340; CAA23627.1; -
DR EMBL: V00341; CAA23629.1; -
DR EMBL: M38303; AAA24583.1; -
DR EMBL: M38304; AAA24580.1; -
DR EMBL: M38292; AAA24579.1; -
DR EMBL: M38293; AAA24581.1; -
DR EMBL: M38287; AAA24585.1; -
DR EMBL: U77436; AAD09605.1; -
DR PIR: A00688; RNECB.
DR PIR: A00689; A00689.
DR SWISS-2DPAGE: P00575; COLI.
DR ECO2DBASE: D157.0; 6TH EDITION.
DR Ecogene: EG10894; rpoB.
DR InterPro: IPR001572; -
DR Pfam: PF00562; RNA_POL_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KM Transference; transcription; DNA-directed RNA polymerase.
FT MTAGEN 813 813 E->K: DISRUPT THE ENZYME'S ACTIVE CENTER.
FT CONFLICT 106 107 ER->G (IN REF. 5).
FT CONFLICT 384 391 LFENLFRS->CSRTCSPT (IN REF. 5).
FT CONFLICT 516 516 D->V (IN REF. 1 AND 9).
SQ SEQUENCE 1342 AA; 150631 MW; F9E95344C54AB118 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
||:|:|:|
440 GEVDDIDHGNR 451

RESULT 8
RPOB_SALTY STANDARD; PRT; 1342 AA.
AC P06173;
DT 01-JAN-1988 (Rel. 06, Created)
DT 13-DEC-1998 (Rel. 37, Last sequence update)
DE BETA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Sverdlov E.D., Listysyn N.A., Guryev S.O., Monastyrskaya G.S.;
RT "Nucleotide sequence of the rpoB gene of Salmonella typhimurium coding
RL for the beta subunit of RNA polymerase.";
RN D01. Biochem. 287:62-65(1986).
RP MEDLINE=89052707; PubMed=3056723;

RA Listysyn N.A., Monastyrskaya G.S., Sverdlov E.D.;
RT "Genes coding for RNA polymerase beta subunit in bacteria.
RT structure/function analysis.";
RL Eur. J. Biochem. 177:363-369(1988).
RN [3]
RP SEQUENCE OF 187-1144 FROM N.A.
RX MEDLINE=86269165; PubMed=3015156;
RA Sverdlov E.D., Listysyn N.A., Guryev S.O., Smitnov V.V.,
RA Gostapshov V.M., Monastyrskaya G.S.;
RT "Genes encoding the beta subunit of bacterial RNA-polymerases. I.
RT Primary structure of the EcoRI-C fragment of the Salmonella
RT typhimurium gene rpoB.";
RL Bioorg. Khim. 12:693-707(1986).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
CC RNA(N).
CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04642; CAA28302.1; -
DR EMBL: M38311; AAA27215.1; -
DR PIR: S01794; RNEBET.
DR StyGene: SG10368; rpoB.
DR InterPro: IPR001572; -
DR Pfam: PF00562; RNA_POL_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KM Transference; transcription; G->A (IN REF. 3).
FT CONFLICT 401 401
SQ SEQUENCE 1342 AA; 150557 MW; EDBF979D3E9E4DB8 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
||:|:|:|
Db 440 GEVDDIDHGNR 451

RESULT 9
RPOB_HAEIN STANDARD; PRT; 1343 AA.
AC P43738;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BETA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z54353; CAA91164.1; -
DR EMBL; AE002371; AAF40591.1; ALT_INIT.
DR TIGR; NMB0132; -
DR InterPro: IPR001572; -
DR Pfam: PF00562; RNA_POL_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Translase; Transcription; DNA-directed RNA polymerase.
FT CONFLICT 2 28 N -> S (IN REF. 1).
FT CONFLICT 28 28 L -> I (IN REF. 1).
FT CONFLICT 224 252 ILIDFYKREFFYSSNGYOTLVADRLKG -> NLGYFLRQ
RNVEFVPRKCSNREGRRSES (IN REF. 1).
FT CONFLICT 284 284 N -> L (IN REF. 1).
FT CONFLICT 350 360 AYISNTLRIDE -> VISPIECVRMK (IN REF. 1).
FT CONFLICT 378 378 MISSING (IN REF. 1).
FT CONFLICT 648 648 A -> G (IN REF. 1).
FT CONFLICT 706 706 A -> P (IN REF. 1).
FT CONFLICT 717 718 VP -> SA (IN REF. 1).
FT CONFLICT 748 749 GG -> A (IN REF. 1).
FT CONFLICT 818 818 F -> L (IN REF. 1).
FT CONFLICT 836 837 GY -> VN (IN REF. 1).
FT CONFLICT 1190 1192 YNG -> SR (IN REF. 1).
FT CONFLICT 1201 1201 A -> S (IN REF. 1).
FT CONFLICT 1244 1246 DDP -> EDA (IN REF. 1).
FT CONFLICT 1373 1373 F -> L (IN REF. 1).
SQ SEQUENCE 1392 AA; 155708 MW; 92AA331A597898F1 CRC64;

Query Match
Best Local Similarity 62.5%; Score 40; DB 1; Length 1392;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
Db 467 GEVDIDHLMNR 478
||| | | | | |
1188 YEAST STANDARD; PRT; 1655 AA.
RESULT 12
1188 YEAST STANDARD; PRT; 1655 AA.
PS2593;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE NUCLEOPORIN NUP188 (NUCLEAR PORE PROTEIN NUP188).
GN NUP188 OR YML103C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN NCB1;
RP SEQUENCE FROM N.A.
RA Zabel U., Doye V., Tekotte H., Wepf R., Hurt E.C.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN NCB1;
RP SEQUENCE FROM N.A.
RA Neirbass U., Rott M.P., Maguire S., Blobel G., Wozniak R.W.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN NCB1;
RP SEQUENCE FROM N.A.
RA STAIN-S288C / AB972;
RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN NCB1;
RP SEQUENCE OF 1187-1205 AND 1611-1629.

```

```

RX MEDLINE=96095775; PubMed=85252578;
RA Aitchison J.D., Rott M.P., Marelli M., Blobel G., Wozniak R.W.;
RT "Two novel related yeast nucleoporins Nup170p and Nup157p:
RT complementation with the vertebrate homologue Nup155p and functional
RT interactions with the yeast nuclear pore-membrane protein Pom152p.";
RL J. Cell Biol. 131:1133-1148(1995).
CC - FUNCTION: INTERACTS WITH THE CORE STRUCTURE OF THE NUCLEAR PORE
CC COMPLEX (NPC). MAY PROVIDE THE NECESSARY ASYMMETRY REQUIRED FOR
CC ANCHORING STRUCTURES SUCH AS CYTOSOLIC FILAMENTS AND THE
CC NUCLEOLAR CAGE.
CC - SUBUNIT: INTERACTS WITH NUP170 AND/OR POM152.
CC - SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X90580; CAA62208.1; -
DR EMBL; U47107; AAA88904.1; -
DR EMBL; X80835; CAA56794.1; -
DR SGD; S0004571; NUP188.
KW Nuclear protein; Transport.
SQ SEQUENCE 1655 AA; 188576 MW; B590051EE60464C CRC64;

Query Match
Best Local Similarity 72.7%; Score 39;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSEIDIETIG 11
Db 1290 GSEIDFIRKNG 1300
||| | | | | |
GUB_RHOMR STANDARD; PRT; 286 AA.
AC P45798;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENDO-BETA-1,3-1,4 GLUCANASE)
DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE).
GN BG1A.
OS Rhodothermus marinus.
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.
OX NCBI_TaxID=29549;
RN NCB1;
RP SEQUENCE FROM N.A.
RA Spillier R., Hreggvidsson G.O., Kristjansson J.K.;
RA Eggertsson G., Palstoft A.;
RT "Cloning and sequencing of a Rhodothermus marinus gene, bg1a, coding
RT for a thermostable beta-glucanase and its expression in Escherichia
RT coli."
RL Eur. J. Biochem. 224:923-930(1994).
CC - FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN
CC BUT NOT ON CMC CELLULOSE OR XYLAN. THE ENZYME HAS A TEMPERATURE
CC OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0.
CC - CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC
CC LINKAGES IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.
CC - SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 004836; AAG60459.1; -
 CC HSSP: P23904; IADR.
 DR InterPro: IPR000757; -
 DR Pfam: PF00722; Glyco_hydro.16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Hydrolyase; Glycosidase; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 286
 FT ACT_SITE 158 158
 FT ACT_SITE 163 163
 FT SEQUENCE 286 AA; 33145 MW; 7215C33624135191 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 286;
 Best Local Similarity 63.6%; Pred. No. 9.1;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SEIDIIETIG 11
 :|||||:1
 156 NGSDIMEHYG 166

RESULT 14

TCPO_CANAL STANDARD; PRT; 540 AA.

AC P47828; 01-FEB-1996 (Rel. 33, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE T-COMPLEX PROTEIN 1, THERA SUBUNIT (TCP-1-THERA) (CCT-THERA).

GN CCT8

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

ON NCBI_Taxid=5476;

RX [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 10231;

RC MEDLINE-96367593; PubMed-8771707;

RA Stoldt V., Rademacher F., Kehren V., Ernst J.F., Sherman F.;

RT *Review: the Cct eukaryotic chaperonin subunits of Saccharomyces

RL cerevisiae and other yeasts.;

RN Yeast 12:523-529(1996).

RP REVISIONS.

RP STRAIN-ATCC 10231;

CC Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON

ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF

ACTIN AND TUBULIN. IN YEAST MAY PLAY A ROLE IN MITOTIC SPINDLE

FORMATION (BY SIMILARITY).

-1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT

FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U37371; AAC31764.1; -

DR InterPro: IPR002194; -

DR InterPro: IPR002423; -

DR Pfam: PF00118; cpn60_TCP1_2.

DR PROSITE: PS00750; TCP1_1; 1.

DR PROSITE: PS00751; TCP1_2; 1.
 DR PROSITE: PS00995; TCP1_3; FALSE_NEG.
 KW Chaperone; ATP-binding; Multigene family.
 SQ SEQUENCE 540 AA; 58918 MW; F962285DA6EB03DE CRC64;

Query Match 60.9%; Score 39; DB 1; Length 540;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GEIDIIET 9
 :|||||
 Db 342 GEIDIIET 349

RESULT 15

CNG_DROME STANDARD; PRT; 665 AA.

AC 024278; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE CYCLIC-NUCLEOTIDE-GATED CATION CHANNEL (CNG CHANNEL).

GN CNG.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

ON NCBI_Taxid=7227;

RX [1]

RP SEQUENCE FROM N.A.

RC MEDLINE-95045396; PubMed-7957070;

RA Baumann A., Frings S., Godde M., Selfert R., Kaupp U.B.;

RT "Primary structure and functional expression of a Drosophila cyclic

nucleotide-gated channel present in eyes and antennae.;"

RL EMBL J. 13:5040-5050(1994).

-1- FUNCTION: APPROXIMATELY 50-FOLD MORE SENSITIVE TO cGMP THAN TO

cAMP. MAY BE INVOLVED IN TRANSDUCTION CASCADES OF BOTH

INTERMEDIATE PHOTORECEPTORS AND OLFACTORY SENSILLAE.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTENNAE AND THE VISUAL SYSTEM.

CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL

FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X89601; CA61760.1; -

DR Flybase; FBgn0014462; Cng.

DR InterPro: IPR000595; -

DR InterPro: IPR002025; -

DR Pfam: PF00914; CNG_membrane; 1.

DR Pfam: PF00027; CNG_binding; 1.

DR PROSITE: PS00888; CNG_BINDING_1; 1.

DR PROSITE: PS00889; CNG_BINDING_2; 1.

DR PROSITE: PS50042; CNG_BINDING_3; 1.

KW Ionic channel; Ion transport; cAMP-binding; Transmembrane.

FT DOMAIN 1 110

FT TRANSSEM 111 131

FT DOMAIN 132 138

FT TRANSSEM 139 159

FT DOMAIN 160 186

FT TRANSSEM 187 207

FT DOMAIN 208 253

FT TRANSSEM 254 274

FT DOMAIN 275 325

FT TRANSSEM 326 346

FT DOMAIN 347 481

FT TRANSSEM 482 502

FT DOMAIN 503 523

FT TRANSSEM 524 544

FT	DOMAIN	503	665	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	437	559	CAMP (BY SIMILARITY).
FT	BINDING	496	496	CAMP (POTENTIAL).
FT	BINDING	511	511	CAMP (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	665 AA;	75922 MM;	9F1BDC5D9581C8DB CRC64;

Query Match 60.9%; Score 39; DB 1; Length 665;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 GEIDIIERTIGNR 13
 ||: :|| |||
 Db 495 GEVSYLEIAGNR 506

Search completed: October 11, 2001, 15:57:05
 Time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:51:42 ; Search time 34.86 seconds
(without alignments)
49.339 Million cell updates/sec

Title: US-09-596-101A-1
Perfect score: 64
Sequence: 1 SGEIDIIETIGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.rodent:*
13: sp.unclassified:*
14: sp.vertebrate:*
15: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	384	5 077072	051333 elsenia foe
2	50	78.1	306	2 051333	051333 cerskovia x
3	47	73.4	361	5 090064	090064 pacifastacu
4	45	70.3	294	2 007242	007242 mycobacteri
5	45	70.3	499	5 026660	026660 strongyloce
6	44	68.8	422	2 09EWR5	09EWR5 streptomyc
7	44	68.8	435	2 068641	068641 cerskovia x
8	43	67.2	285	2 09F3A0	09F3A0 streptomyc
9	43	67.2	1383	2 09KJMS	09KJMS bartonella
10	43	67.2	1383	2 09KJG4	09KJG4 bartonella
11	41	64.1	736	10 09SVX7	09SVX7 arabidopsis
12	40	62.5	342	2 09F406	09F406 neisseria m
13	40	62.5	342	2 09F405	09F405 neisseria m
14	40	62.5	342	2 09F404	09F404 neisseria m
15	40	62.5	342	2 09E075	09E075 neisseria m
16	40	62.5	342	2 09ERT6	09ERT6 neisseria m
17	40	62.5	342	2 09ERT6	09ERT6 neisseria m
18	40	62.5	395	5 017492	017492 anopheles g
19	40	62.5	806	11 09WVF4	09WVF4 mus musculu

20	40	62.5	1342	2 09L9J8	09L9J8 salmonella
21	40	62.5	1375	2 09KJ30	09KJ30 vibrio chol
22	39	60.9	276	2 052734	052734 rhodothermu
23	39	60.9	642	2 09WXX1	09WXX1 thermotoga
24	39	60.9	646	2 060039	060039 thermotoga
25	39	60.9	665	5 09V7L5	09V7L5 drosophila
26	39	60.9	967	10 09L4V7	09L4V7 arabidopsis
27	39	60.9	1324	2 059328	059328 clostridium
28	38	59.4	263	2 09WZC8	09WZC8 thermotoga
29	38	59.4	297	1 073951	073951 pyrococcus
30	38	59.4	298	2 P73049	P73049 synecocyst
31	38	59.4	544	2 050223	050223 thibacillus
32	38	59.4	608	14 09YV22	09YV22 melanoplus
33	37	57.8	316	10 09W2H6	09W2H6 arabidopsis
34	37	57.8	664	2 09K9Z0	09K9Z0 bacillus ha
35	37	57.8	794	2 09R1A3	09R1A3 streptomyc
36	37	57.8	877	2 045095	045095 bacillus ci
37	37	57.8	908	1 028789	028789 archaeoglob
38	37	57.8	1292	5 09Y0C8	09Y0C8 drosophila
39	37	57.8	1404	5 045251	045251 caenorhabdi
40	37	57.8	1429	5 015764	015764 tetrahymena
41	36.5	57.0	188	1 09Y9P2	09Y9P2 aeropyrum p
42	36	56.2	88	1 09HS27	09HS27 halobacteri
43	36	56.2	223	1 027105	027105 methanobact
44	36	56.2	310	2 09KNJ8	09KNJ8 vibrio chol
45	36	56.2	352	5 09V4R3	09V4R3 drosophila

ALIGNMENTS

RESULT 1
077072 PRELIMINARY; PRT; 384 AA.
ID 077072;
AC 077072;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE COELOMIC CYTOLYTIC FACTOR 1.
GN CCF1.
OS Eisenia foetida (Common branding worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RX MEDLINE=98406152; PubMed=9733802;
RA Beschyn A., Bilej M., Hanssens F., Raymakers J., Van Dyck E.,
RA Reyets H., Blys L., Gomez J., De Baetselier P., Timmermans M.,
RT Identification and cloning of a glucan- and lipopolysaccharide-
RT binding protein from Eisenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade.*;
RL J. Biol. Chem. 273:24948-24954(1998).
DR EMBL: AF030028; AAC35887.1;
SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

Query Match 100.0%; Score 64; DB 5; Length 384;

Best Local Similarity 100.0%; Pred. No. 0.0028;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
Db 178 SGEIDIIETIGNR 190

RESULT 2
ID 051333 PRELIMINARY; PRT; 306 AA.
AC 051333;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE BETA-1,3-GLUCANASE IIA.
 GN BGLIIA.
 OS Oerskovia xanthineolytica.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micrococcales; Cellulomonadaceae; Cellulomonas.
 NCBI_TaxID=1826;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL 6109.
 RX MEDLINE-96345651; PubMed-8755914;
 RA Ferrer P., Halkier T., Hedegaard L., Savva D., Diers I., Asenjo J.A.;
 RT "Nucleotide sequence of a beta-1,3-glucanase isoenzyme IIA gene of
 RT Oerskovia xanthineolytica LL 6109 (Cellulomonas cellulans) and initial
 RT characterization of the recombinant enzyme expressed in Bacillus".
 RT subills."
 RL J. Bacteriol. 178:4751-4757(1996).
 DR EMBL: U56935; AAC44371.1; -.
 DR HSSP: P23904; IAJK.
 DR InterPro: IPR000757; -.
 DR Pfam: PF00722; Glyco_hydro.16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 SEQUENCE 306 AA; 32835 MW; CDBDB8C1A6F8DC04 CRC64;

Query Match 78.1%; Score 50; DB 2; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.69;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 180 SGEIDIMENVGN 191

RESULT 3
 ID 090064 PRELIMINARY; PRT; 361 AA.
 AC 090064;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE LIPOPOLYSACCHARIDE AND BETA-1,3-GLUCAN BINDING PROTEIN PRECURSOR.
 GN LGBP.
 OS Pacifastacus leniusculus (Signal crayfish).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Astacoidea; Astacidae; Pacifastacus.
 NCBI_TaxID=6720;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HEMOCYTES;
 MEDLINE-20092910; PubMed-10625682;
 RA Lee S., Wang R., Soderhall K.;
 RT "A lipopolysaccharide- and beta-1,3-glucan-binding protein from
 RT hemocytes of the freshwater crayfish Pacifastacus leniusculus:
 RT purification, characterization, and cDNA cloning.";
 RL J. Biol. Chem. 275:1337-1343(2000).
 DR EMBL: AJ250128; CAB55353.1; -.
 KW Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 361 LIPOPOLYSACCHARIDE AND BETA-1,3-GLUCAN
 FT BINDING PROTEIN.
 SO SEQUENCE 361 AA; 41043 MW; 93A5E67911ED6619 CRC64;

Query Match 73.4%; Score 47; DB 5; Length 361;
 Best Local Similarity 75.0%; Pred. No. 2.9;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 172 SGEIDIVESRGN 183

RESULT 4
 ID 007242 PRELIMINARY; PRT; 294 AA.
 AC 007242;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROBABLE BETA-1,3-GLUCANASE.
 GN RV0315 OR MTCY63.20.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-96295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: Z96800; CAB09586.1; -.
 DR TuberculList: RV0315; -.
 DR InterPro: IPR000757; -.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN.1.
 SEQUENCE 294 AA; 32186 MW; 7B9897BA6740BEB0 CRC64;

Query Match 70.3%; Score 45; DB 2; Length 294;
 Best Local Similarity 75.0%; Pred. No. 5.1;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 169 SGEIDLEWIGN 180

RESULT 5
 ID 026660 PRELIMINARY; PRT; 499 AA.
 AC 026660;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE BETA 1,3-GLUCANASE.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 NCBI_TaxID=7668;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96270625; PubMed-8692900;
 RA Bachman E.S., McClay D.R.;
 RT "Molecular cloning of the first metazoan beta-1,3 glucanase from eggs
 RT of the sea urchin Strongylocentrotus purpuratus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6808-6813(1996).
 DR EMBL: U49711; AAC47233.1; -.
 SO SEQUENCE 499 AA; 55274 MW; D863F36E0CF5AC CRC64;

Query Match 70.3%; Score 45; DB 5; Length 499;
 Best Local Similarity 66.7%; Pred. No. 9.5;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12

DB 326 SGEIDLVESRGN 337

|||||:|:|

RESULT 6

09EMR5 PRELIMINARY; PRT; 422 AA.

AC 09EMR5; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

GN PUTATIVE SECRETED HYDROLASE.

3SC60.19.

Streptomyces coelicolor.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

NCBI_TaxID=1902;

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.; Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;

Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

MEDLINE=97000351; PubMed=8843436;

Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,

Kinashi H., Hopwood D.A.;

"A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).

EMBL: AL45945; CAC14352.1; -

Hydrolase.

SEQUENCE 422 AA; 44143 MW; 7105047BE26F3020 CRC64;

SO

Query Match 68.8%; Score 44; DB 2; Length 422;

Best Local Similarity 72.7%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11

|||||:|:|

170 SGEIDMENVG 180

RESULT 7

068641

AC 068641; PRELIMINARY; PRT; 435 AA.

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE BETA-1,3-GLUCANASE II.

BGLI.

Oerakova xanthineolytica.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Micrococccineae; Cellulomonadaceae; Cellulomonas.

NCBI_TaxID=1826;

SEQUENCE FROM N.A.

STRAIN-LL G109;

Ventom A.M., Asenjo J.A.;

Enzyme Microb. Technol. 13:71-75(1991).

SEQUENCE FROM N.A.

STRAIN-LL G109;

MEDLINE=96409238; PubMed=8814220;

Parrado J., Escuredo P.R., Conejero-Lara F., Kotik M., Ponting C.P.,

Hydrolase.

SEQUENCE 285 AA; 30479 MW; 6977A032A1A1451 CRC64;

SO

Asenjo J.A., Dobson C.M.; "Molecular characterisation of a thermostable beta-1,3-glucanase from

Oerakova xanthineolytica.";

Biochim. Biophys. Acta 1296:145-151(1996).

SEQUENCE FROM N.A.

STRAIN-LL G109;

MEDLINE=96250169; PubMed=8659924;

Ferrer P., Hedegaard L., Halkier T., Diers I., Savva D., Asenjo J.A.;

"Molecular cloning of a lytic beta-1,3-glucanase gene from Oerakova

xanthineolytica ligG109. A beta-1,3-glucanase able to selectively

penetrate the yeast cell wall.";

Ann. N. Y. Acad. Sci. 782:555-566(1996).

SEQUENCE FROM N.A.

STRAIN-LL G109;

Ferrer P., Andrews B.A., Asenjo J.A., Hedegaard L., Diers I.;

Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

EMBL: AF052745; AAC38290.1; -

HSSP; P23904; IABK.

InterPro: IPR000757; -

InterPro: IPR000772; -

Pfam: PF00652; Rich. B. lectin; 1.

Pfam: PF00722; Glyco_hydro_16; 1.

PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.

SMART: SM00458; RICIN; 1.

SEQUENCE 435 AA; 46097 MW; 00F087BE64C0F58 CRC64;

Query Match

Best Local Similarity 68.8%; Score 44; DB 2; Length 435;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11

|||||:|:|

169 SGEIDMENVG 179

DB

RESULT 8

09F3A0

AC 09F3A0; PRELIMINARY; PRT; 285 AA.

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE PUTATIVE SECRETED HYDROLASE.

SC5F1.29C.

Streptomyces coelicolor.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

NCBI_TaxID=1902;

SEQUENCE FROM N.A.

STRAIN-A3(2);

Saunders D.C., Harris D.; Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;

Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

MEDLINE=97000351; PubMed=8843436;

Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,

Kinashi H., Hopwood D.A.;

"A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).

EMBL: AL450165; CAC16455.1; -

Hydrolase.

SEQUENCE 285 AA; 30479 MW; 6977A032A1A1451 CRC64;

SO

Query Match 67.2%; Score 43; DB 2; Length 285;

Best Local Similarity 63.6%; Pred. No. 11; Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGR 11
|||||:|:|:|
DB 160 SGEIDVMEVNG 170

RESULT 9

O9KJMS PRELIMINARY; PRT; 1383 AA.

AC O9KJMS: 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE RNA POLYMERASE BETA SUBUNIT.

GN RPOB.
Bartonella quintana (Rochalimaea quintana).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Bartonellaceae; Bartonella.

OX NCB1_Taxid=803;

RP SEQUENCE FROM N.A.

RC STRAIN-FULLER.

RA Renesto P., Gouvernet J., Drancourt M., Raoult D.;

RT "New approach to identifying Bartonella species on the basis of the

rtprob gene sequencing.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

SUBSTRATES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +

RNA(N) (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

DR EMBL; AF165994; AAF80850.1; -

DR InterPro: IPR001572; -

DR Pfam: PF00562; RNA_POL_B; 1.

DR PROSITE; PS01166; RNA_POL_BETA; 1.

KW DNA-directed RNA polymerase; Transcription; Transferase.

SO SEQUENCE 1383 AA; 154780 MW; 9B007A67C872498E CRC64;

Query Match 67.2%; Score 43; DB 2; Length 1383;

Best Local Similarity 66.7%; Pred. No. 71; Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

2 GEIDIIETIGR 13

|||||:|:|:|

DB 454 GEIDIDIDNIGR 465

RESULT 10

O9KJG4 PRELIMINARY; PRT; 1383 AA.

AC O9KJG4: 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE RNA POLYMERASE BETA SUBUNIT.

GN RPOB.

OS Bartonella henselae (Rochalimaea henselae).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Bartonellaceae; Bartonella.

OX NCB1_Taxid=38333;

RP SEQUENCE FROM N.A.

RC STRAIN-HOUSTON-1;

RA Renesto P., Gouvernet J., Drancourt M., Raoult D.;

RT "New approach to identifying Bartonella species on the basis of the

rtprob gene sequencing.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

SUBSTRATES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +

RNA(N) (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

DR EMBL; AF171070; AAF87049.1; -

DR InterPro: IPR001572; -

DR Pfam; PF00562; RNA_POL_B; 1.

DR PROSITE; PS01166; RNA_POL_BETA; 1.

KW DNA-directed RNA polymerase; Transcription; Transferase.

SO SEQUENCE 1383 AA; 154852 MW; 2CE87E06A3207B6 CRC64;

Query Match 67.2%; Score 43; DB 2; Length 1383;

Best Local Similarity 66.7%; Pred. No. 71; Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGR 13

|||||:|:|:|

DB 454 GEIDIDIDNIGR 465

RESULT 11

O9SVX7 PRELIMINARY; PRT; 736 AA.

AC O9SVX7: 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DE HYPOTHEtical 84.2 KDA PROTEIN.

GN F15B8.180.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCB1_Taxid=3702;

RP SEQUENCE FROM N.A.

RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W.,

RA Mayer K.F.X., Lemcke K., Schueller C., Quetier F., Salanoubat M.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL049660; CAB41192.1; -

DR InterPro: IPR000561; -

DR Pfam; PF000249; -

DR PRINTS; PR00011; EGFAMININ.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.

DR PROSITE; PS01186; EGF_2; 2.

KW EGF-like domain; Glycoprotein; Hypothetical protein.

SO SEQUENCE 736 AA; 84202 MW; 349E0F1E6A28C9A CRC64;

Query Match 64.1%; Score 41; DB 10; Length 736;

Best Local Similarity 72.7%; Pred. No. 77; Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 EIDIIETIGR 13

|||||:|:|:|

DB 109 EVDIIETIGR 119

RESULT 12

O9F406 PRELIMINARY; PRT; 342 AA.

AC O9F406: 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270494; CAC14798.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37237 MW; E4C6B3719348B0AB CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 13

O9F405 PRELIMINARY; PRT; 342 AA.
 AC O9F405;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.12;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270501; CAC14805.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37268 MW; 5C08CDA6C25AFBCF CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 14

O9F404 PRELIMINARY; PRT; 342 AA.
 AC O9F404;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 SQ

GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270502; CAC14806.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37264 MW; D04BAA7653E5B324 CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 15

O9EU75 PRELIMINARY; PRT; 342 AA.
 AC O9EU75;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:14:PL.12, C:2B:PL.2, C:2B, AND:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270504; CAC14808.1; -.
 DR EMBL: AJ270500; CAC14804.1; -.
 DR EMBL: AJ270503; CAC14807.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37253 MW; 2CDB9BD9709A5DEDE CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

Search completed: October 11, 2001, 15:56:43
 Job time: 301 sec

DR WPI: 1999-385905/32.
DR N-PSDB: AAX83611.

XX Eiseinia foetida polypeptides derived from coelomic cytolytic factor 1
XX Claim 2; Page 48-49; 49pp; English.

CC The present sequence represents the Eiseinia foetida coelomic cytolytic
CC factor 1 (CCF-1). The protein has antiparasitic, antibacterial and
CC anti-inflammatory activity. Recombinant coelomic cytolytic factor 1
CC (rCCF-1) is trypanolytic for the African trypanosome Trypanosoma brucei
CC in a dose-dependent manner. The trypanolytic activity of rCCF-1 can be
CC inhibited by anti-CCF-1 and anti-tumour necrosis factor (TNF)/TNF-
CC monoclonal antibodies. Furthermore, N,N'-diacetylethanolamine inhibits
CC potentially trypanolytic activity of rCCF-1. These data corroborate the
CC findings that CCF-1 shares a trypanolytic, lectin-like domain with
CC TNF-alpha. CCF-1 is useful to treat trypanosomal or bacterial infections
CC or cancer. The proteins and peptides are also useful in tumour therapy,
CC inflammation and other areas of immunology. The annexed peptide is
CC derived from a 4.42 kDa cytolytic protein named CCF-1 that binds
CC lipopolysaccharide and beta-1,3-glucan. The factor resembles the
CC vertebrate tumour necrosis factor-alpha (TNF-alpha), and may be used as
CC an alternative for TNF-alpha.

XX Sequence 384 AA;

Query Match 100.0%; Score 2240; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-204; Indels 0; Gaps 0;
Matches 384; Conservative 0; Mismatches 0;

OY 1 MHWTLVLLCLLGEFGAFYDMDQYHIVWDEDFYFGAKWQHEVTAVGSGSEFQATYTOP 60
DB 1 mhwtlvllcllgefgafydwqyhyhwdeidyfgakwqhevtavgsnsefqltqtd 60
OY 61 GANSFYRQDKLFIKPTLADNINPQTGAPEGIDPMYNGVLDVMAATGACTTNDNNGCYRT 120
DB 61 gansfyrdgklfipkptladninpqtgapegidpmyngvldvmaatgacttndnngcyrt 120
OY 121 GAAGDTPPMASRVTRFQKYSFTHGKVVVNAKMPVGDMLPAILMIPEDWVYGGMPISGE 180
DB 121 gaagdtppmasrvtrfqkysfthgkvvvnaakmpvgdmlpailmipedwvyggmpisge 180
OY 181 IDIIETIGNRDEKNTGGEFGIOLKMGSTMHGPGMDNRYTLTSLPKHSDMNYGDNFHT 240
DB 181 idiietignrdekntggefgiolkmgstmhgpgmdnrytltslpkhsdmnygdnfht 240
OY 241 FWFDMSPNGLRFFVDDENQALDVPYPLIDANFWMVDFWEGKFWLPQYENDNPWAGTIN 300
DB 241 fwfdmspnqlrffvddengalldvpyplidanfwmvdfwewgkfwlpqyendnpwagtn 300
OY 301 LAPFOQNEFIILNVAVGTNGFIPDCINRGSPALQKPKWSNDWYNDAKRKFFDAGNW 360
DB 301 lapfoqnefiilnvavgtngfipdcinrgspalqkpkwsndwynyndamrkffidargnw 360
OY 361 KWTWDEGDNNAWQVYIRYTKRN 384
DB 361 kwtwdegdnnaqwvirytkrn 384

RESULT 2
AAM29457
ID AAM29457 standard; Protein; 303 AA.

XX AAM29457;

XX 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica beta-1,3-glucanase.

KW Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
KV fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LW6109 (DSM 10297).

XX Key Location/Qualifiers
FH Peptide 1..52
FT /label= sig_peptide
FT Protein 53..303
FT /label= Mat_protein

XX MO9739114-A1.

XX 23-OCT-1997.

XX 14-APR-1997; 97WO-DK00160.

XX 23-AUG-1996; 96DK-0000885.

XX 12-APR-1996; 96DK-0000427.

XX (NOVO) NOVO-NORDISK AS.

XX Diers I, Ferrer P, Halkier T, Hedegard L;

DR WPI: 1997-526451/48.

DR N-PSDB: AAT89157.

XX New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
XX xanthineolytica, used particularly for the lysis of microbial cells
XX for obtaining desirable products

XX Example 2; Page 42-43; 64pp; English.

CC This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (BG) activity. Its amino acid sequence was deduced from an
CC isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs
CC that encode the novel BG (see also AAM29455), a mannose binding
CC domain (see AAM29458) or a full-length enzyme, i.e. BG with mannose
CC binding domain (see AAM29456), can be used to produce recombinant BG
CC polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. BG polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.

XX Sequence 303 AA;

Query Match 15.6%; Score 350; DB 18; Length 303;
Best Local Similarity 28.5%; Pred. No. 3.1e-25;
Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

OY 26 IWQDEDFYFDG-----AKWQHEVTATGSGSEFQATYQDGAASFVRDGLFIKPTLAD 80
DB 57 llwdefdgaagsapnrvavnhetgahgwnaqlnylastranaal-dgq----- 105
OY 81 NINPQTGAPEGIDPMYNGVLDVMAATGACTTNDNNGCYRTGAAGDIPPMASRVTRFOKY 140
DB 106 -----gnlvlt-----regdgsy-----tsamltgqky 131
OY 141 SFTHGRVYVNAKMPVGDMLPAILMIPEDWYGGMPRSGEIDIIETIGNRQFKNTGSEFL 200
DB 132 qpqygrilearilqprgqivpatwmlggsifpqtppssgeidienyvfephrvhg----- 187
OY 201 GIQKMGSTMHGPGMDNRYTLTSLPKHSDMNYGDNFHTFWMSPGRLFFVDDENQA 260
DB 188 -----tvh-gpysggs-gltgmyqhpqgwsfadthtflavdkkpgelctwfd--gqg 236
OY 261 LLDVYPLIDANFWMVDFWEGKFWLPQYENDNPWAGTINLAPDQNEFIILNVAVGTN 320
DB 261 lladvpyplidanfwmvdfwewgkfwlpqyendnpwagtnlapdqnepiilnvavgttn 320

Db 397 aaggttcf.rdgsltsqg---vfkpw-----rdsarkaaevhfwrlnmsdwfrpsqp--- 443
Oy 371 NAMQVDYIKV 380
::||::|
Db 444 -slvdtkrv 452

RESULT 6
AAR89137
ID AAR89137 standard; Protein; 467 AA.
XX
AC AAR89137;
XX
DT 22-AUG-1996 (first entry)
XX
DE Bombyx mori full length LPS-binding protein.
XX
RV 1lipopolysaccharide binding protein; Bombyx mori; haemolymph; primer;
XX Enterobacter cloacae; PCR; polymerase chain reaction; amplification;
XX probe; LPS-complex; septic shock; infection; transgenic plant; vine;
XX tobacco; tomato; potato; fungal infection; fungus.

XX
OS Bombyx mori.
XX
FH Key Location/Qualifiers
FT Peptide 1..12
FT Peptide /note= "signal peptide"
FT Peptide 13..467
FT Modified-site /note= "mature protein"
FT 182
FT /note= "putative N-glycosylation site"

XX
PM FR2721032-AL.
XX
XD 15-DEC-1995.
XX
PD 09-JUN-1994; 94FR-0007083.
XX
PE 09-JUN-1994; 94FR-0007083.
XX
PR 09-JUN-1994; 94FR-0007083.
XX
PA (INSP) INST PASTEUR.
PI Brey PT, Lee W;
PI XX
PI XX
DR MPI: 1996-060094/07.
DR N-PSDB; AAT10279.
XX

PT New protein from Bombyx mori that binds bacterial
PT lipopoly-saccharide - esp. used to treat septic shock, also DNA
PT encoding it, for producing transgenic plant(s) resistant to fungal
PT attack

PS Claim 5; Page 42-43; 53pp; French.
XX
XX

CC This is the amino acid sequence of the full length lipopolysaccharide
CC (LPS) binding protein from Bombyx mori. The protein was isolated from
CC the haemolymph of fifth stage B.mori larvae after injection with
CC Enterobacter cloacae strain 57-9. Partial amino acid sequence was used
CC to generate PCR primers (AAT10280-1). These amplified a fragment of the
CC gene used as a probe to obtain the full length gene by screening a cDNA
CC library derived from B.mori previously injected with heat-killed
CC E.coliaceae. The protein or fragments of it, can be used to treat diseases
CC associated with LPS complexes e.g. septic shock, to remove LPS from
CC products used for injection and to protect transgenic plants e.g. vines,
CC tobacco, tomato or potato, against fungal infections.
XX
SQ Sequence 467 AA;

```

QY 27 WMOEDYEDFCAGAKNO-HEVTAITGGGNSSEPOLYTQDDGANSFVROSKLEIKPFLTLADNINPQ 85
Db 147 ifeeqfsgldenvwqleqlyirlyihpeyrfvsyqnnlvtstadsnllh-----naklj 1399
QY 86 TGAP-FGHDPMYNGVLADWAMYAGACTNTDNGGCRRTGAAGDI-PPASARVRFQKXSF 143
Db 200 qhmpgfiddslysgtlh---lfsqctas-aacalcqsgsdellpivrsygl-tsigrft 254
QY 144 HGRVVVHAKMFVGDMLPMALTMLEPDWNYGGP-RSGEIDIIEITGN-----RDKNT 135
Db 255 ygtveiraklpqgdwlypellleflfkygsmuysasvkvkacargnaelysgpndysnt 314
QY 196 GSEFLGDKMGSTNMHWGMD-DNRMYLTSLRKHSDDMNNGDNHRTFPMSPGGLFEV 254
Db 315 -----vlygppimldecentfisktkrrtdgtswqdsfhlcsyvgqtpcfofials 352
QY 255 DDENQALDVEYPLIDANPMWVWDWEMGKPMPLPOLYENDNPMAGGTNLAPFDONFHTLNV 314
Db 369 dgeeararvearpdael-----pvcchaphrllqsgsqmqrtdhnyltly 408
QY 315 AVGTNGTIPPGCINRGSDPLQKPMWSNGDWTNDAMRK----FPDANGNKWTYDDEGDN 370
Db 409 aagstief-rdgsiltsisg---vtkpw-----idsarkasvhtfrlmsdswfwrsgp--- 455
QY 371 NAMQVDYIRV 380
Db 456 -slvdltkv 464

```

XX	RESULT	7
XX	AA11599	
ID	AA11599 standard; Protein; 261 AA.	
XX		
AC	AA11599;	
XX		
DT	18-JUN-1991 (first entry)	
XX		
DE	Beta-1,3-glucanase.	
XX		
KW	Alkalophilic; heat resistant.	
XX		
OS	Bacillus sp.	
XX		
PN	JP03053883-A.	
XX		
PD	07-MAR-1991.	
XX		
PF	20-JUL-1989; 89JP-0185928.	
XX		
PR	20-JUL-1989; 89JP-0185928.	
XX		
PA	(SHKJ) SHINGIJUTSU KAIHAISU.	
XX		
XX		
DR	WPI; 1991-113290/16.	
DR	N-PEDB; AAQ11293.	
PT	Heat resistant beta-1,3-glucanase gene DNA - derived from	
XX	alkali-compatible Bacillus sp.	
PS		
XX	Claim 7; Fig 2; 11pp; Japanese.	
CC		
CC	The new enzyme is derived from alkalophilic Bacillus Sp. The	
CC	enzyme acts on beta-1,3-glucans to form glucose and lamnaribiose	
CC	It is stable over a wide pH range with opt. activity at weakly	
CC	acidic pH and has good heat resistance.	
XX		
SQ	Sequence 261 AA;	

Query Match	14.0%;	Score 314;	DB 17;	Length 467;
Best Local Similarity	27.3%;	Pred. NO. 1.4e-21;		
Matches 101;	Conservative 58;	Mismatches 143;	Indels 68;	Gaps 17;

Query Match 13.8%; Score 310; DB 12; Length 261;
Best Local Similarity 26.9%; Pred. No. 1.6e-21;
Matches 100; Conservative 35; Mismatches 83; Indels 154; Gaps 17;

[illegible]

XX	RESULT	8
XX	AA97362	
XX	ID	AA97362 standard; Protein: 306 AA.
XX	AC	AA97362;
XX	DT	03-OCT-1996 (first entry)
XX	DE	Oerskovia beta-1,3-glucanase.
XX	KM	Beta-1,3-glucanase; Cellulomonas cellulans; Bacillus subtilis;
XX	KM	lytic enzyme; beta-glucan degradation; cell wall lysis;
XX	KM	pigment; colorant; flavour; yeast extract; protoplast.
XX	OS	Oerskovia xanthineolytica strain LMG109 (DSM 10297).
XX	FT	Key
XX	FT	Peptide
XX	FT	Peptide
XX	FT	Protein
XX	FT	Location/Qualifiers
XX	PN	1..35
XX	PD	/label- Sig_peptide
XX	PD	36..63
XX	PF	/label- Pro_peptide
XX	PF	64..306
XX	PR	/label- Mat_protein
XX	PA	MO9612013-A1.
XX	PI	25-APR-1996.
XX	PI	16-OCT-1995;
XX	PI	95WO-DK00414.
XX	PI	14-OCT-1994;
XX	PI	94DK-0001192.
XX	PI	(NOVO) NOVO-NORDISK AS.
XX	PI	Asenjo JA, Diers I, Rerrer P, Halkier T, Hedegaard L;
XX	PI	Savaa D;
XX	PI	WPI: 1996-222000/22.
XX	PI	N-PSDB; AAT29043.
XX	PI	DNA construct encoding enzyme with beta-1,3-glucanase activity -
XX	PI	useful for modifying or degrading beta-glucan contg. material and in
XX	PI	the prepn. of e.g. food colourants, flavourings and yeast extracts

ps Claim 1, Page 42-43; 60pp; English.
 xx
 xx A novel beta-1,3-glucanase (AAR97362) from Oerskovia xanthineolytica
 CC L16109 is useful for degrading or modifying beta-glucan-contg.
 CC material. Its amino acid sequence was deduced from a genomic
 CC DNA sequence (AAR29943) isolated from an O. xanthineolytica library.
 CC Recombinant beta-1,3-glucanase can be produced on a large scale
 CC using transformed host cells, esp. Bacillus subtilis DNB85 or
 CC ToC46. Protease-free beta-1,3-glucanase can be obd. that is useful
 CC for lysing fungal cell walls, allowing recovery of intracellular
 CC proteins. The enzyme is also useful for the prep. of protoplasts
 CC and for the prodn. of pigments, colorants, flavours, yeast extract
 CC and pharmaceuticals.
 xx
 xx Sequence 306 AA;
 SQ

Query Match	Similarity	13.2%	Score 296	DB 17	Length 306
Best Local	93.1	Conservative	36	Mismatches 99	Indels 134; Gaps 13
26	IWMDEPFYFDGAK-----	WHEVATATGGGSEFQLYTQDGAASFVRDGLRIFKPTLLAD	80		
DB	68 lawsefdgaagsapndvwhhetgagysanaelqytltsrnsal-dgq-----		116		
QY	81 NINPOTGARFGEDPMYKNCVLDMVAMATYAGACTNDDNCCYRTGAAGDIPRAMSARVTFQKY	140			
DB	117 -----gnlyvtal-----qesdgsy-----tsarlitqgnv	142			
QY	141 SEFHGRVYVYHAKMPYGDMLMPALIMLPEDMVYGGWPRSGEIDIIETIGNRDKPNTGEPFL	200			
DB	143 qpgfgirleairldiprggqfwsafmwnvanrlptprrtsgeidimenvgnaphvbg----	198			
QY	201 GLOKMGSTMHMGPCMD--NKRKYLITSLPRKSDQWNTGDNHTHTWPFMSPGCLFEFFDDENO	259			
DB	199 -----lvh-pgpysgdn--qimtgtyqhpqgwsaddlmfigidwcpelctlvqd--gq	246			
QY	260 ALLDVPYPLIDNPMMWVYFEMWEGKFWLPOYENDNPFWAGSTNLAPFDONFHTLMAVAGST	319			
DB	247 eyhrvtladvgngv-----vfdqpflllnvalsg-	277			
QY	320 NKFIFDGCILNRGGDALAKPWSNGDMYNDAMKFFDARKNMKWTWDEDDNNAMQDYIR	379			
DB	278 -----qwpqndatpfp-----qmkvdyvr	299			
QY	380 yy 381				
DB	300 YV 301				
RESULT	9				
AAM56275	AAM56275 standard; Protein: 422 AA.				
XX	AAM56275;				
XX	18-AUG-1998 (first entry)				
XX	Flavobacterium keratolyticus endo-beta-galactosidase (ENDO-A) protein.				
XX	alpha-N-acetylgalactosaminidase; endo-beta-galactosidase; ENDO-A;				
XX	erythrocyte; transfusion; blood typing reaction; keratan sulphate;				
XX	A antigen.				
XX	Flavobacterium keratolyticus.				
OS	Key				
FH	Location/Qualifiers				
FT	1..46				
FT	/note= "Signal peptide"				
FT	47..422				
FT	/note= "F. keratolyticus ENDO-A protein"				
XX	W09611246-A2.				

XX 19-MAR-1998.
 PD 04-SEP-1997; 97WO-0515532.
 PF 11-SEP-1996; 96US-0712072.
 PR (NYBL-) NEW YORK BLOOD CENTER INC.
 PA Goldstein J, Hurst R, Leng L, Lenny L, Zhu A;
 PI WPI: 1998-207405/18.
 DR N-PSDB; AAV22722.
 XX New isolated endo-beta-galactosidase - from Flavobacterium
 PT keratolyticus, used particularly for de-antigenising human
 XX erythrocytes bearing A1 antigen for blood transfusions.
 Claim 6; Fig 2A-2B; 45pp; English.
 The present sequence represents the Flavobacterium keratolyticus
 endo-beta-galactosidase (designated ENDO-A) protein. The ENDO-A
 protein can be used in conjugation with alpha-N-acetylgalactosaminidase
 to remove externally and internally linked A antigen on group A1
 erythrocytes. The resulting erythrocytes may be transfused into
 individuals who would be otherwise unable to safely tolerate a
 transfusion of type A1 blood. The ENDO-A protein can also be used to
 degrade keratan sulphate (e.g. in the food industry), or in blood typing
 reactions.
 SO Sequence 422 AA;
 Query Match 12.8%; Score 286.5; DB 19; Length 422;
 Best Local Similarity 27.0%; Pred. No. 5.1e-19;
 Matches 103; Conservative 33; Mismatches 100; Indels 147; Gaps 15;
 QY 10 LIFGEGAFPTDMQYHIVWODEPDIY---FDGAKMOHEVYATGCGN---SEFQLY-TODGA 62
 DB 45 llnatvatld---yellwdefnsgsfidstkwysadrgtvaawkymtsipaysqds 101
 QY 63 NSFYRQDKLFKPTLLADNINPQTGAPGCTDFMTNGVLDVWAMGACNTNNGCYRTGA 122
 DB 102 nlvlt-----mhnayv----- 112
 QY 123 AGDIPPMASARVTFQKYSFTHGKRVVYAKMPVGDMLPAILMPLPED-WYVGMPRSGEI 181
 DB 113 agdpvayhaggyvsmgkfsmtgykvevraktfgtgrspatammpepataygypscgei 172
 QY 182 DIETIGNRDPFKNTGGEFLGIQKMGSTWMMGPMGMDNRYWLTSLPKHSDNNYGDNEHTE 241
 DB 173 dsmehvneev-----myhtlhngsvlnaag---gstaskaatyntcd-ytly 216
 QY 242 WFMSPGALFEFVYDENQALLDVPYPLIDANPMWVDFEWGKPMPLPOYENDNPMAGNTL 301
 DB 217 tmlwspndltfyt---nnsf-----qylvaivsqqggtq 247
 QY 302 APEDQNFHILNAVAG-GTNGFIPDGCINRGDPALQKPMGNGDMYNDAMKRFEDARGNW 360
 DB 248 wplvtpfylllmgsggsgwpaatnadi-----pfs----- 278
 QY 361 KWTWDEGDNNAQVDYIRYTK 382
 DB 279 -----mqvdyvryvyk 288

XX (1-3)-beta-D-glucan sensitive factor.
 DE (1-3)-beta-D-glucan sensitive factor; antifungal agent;
 XX (1-3)-beta-D-glucan sensitive factor; antifungal agent;
 KW mycosis diagnosis.
 XX Limulus sp.
 OS Key Location/Qualifiers
 FH Peptide 1..19
 FT Label- sig_peptide
 XX W09501432-A.
 XX 12-JAN-1995.
 PD 29-JUN-1994; 94WO-JP01057.
 PF 29-JUN-1993; 93JP-0184403.
 PR (SEBK) SEIKAGAKU KOGYO CO LTD.
 PA Iwanaga S, Muta T, Oda T, Seki N;
 PI WPI: 1995-060996/08.
 DR N-PSDB; AA081334.
 DR DNA encoding a polypeptide comprising a tetrapeptide motif at
 PT least once - which may be used as an antibacterial and
 PT antifungal.
 XX Claim 9; Pages 23-32; 51pp; Japanese.
 PS AA081334 encodes AAR67915 a (1-3)-beta-D-glucan sensitive factor, it
 CC has a high affinity for the (1-3)-beta-D-glucan found in fungal
 CC cell walls. The protein is therefore useful for clinically
 CC diagnosing mycosis, and as an antifungal agent for the removal
 CC of fungi.
 SO Sequence 673 AA;
 Query Match 9.3%; Score 209; DB 16; Length 673;
 Best Local Similarity 21.8%; Pred. No. 2.2e-11;
 Matches 85; Conservative 49; Mismatches 106; Indels 150; Gaps 15;
 QY 5 LVVL-CLIFGEGFA---PTDMQYHIVWODEPDIYDGAQMOHEV---TATGGNSEFQLYT 58
 DB 2 lvlccvllhvgvalccshcpkqglwvdsdefltnglssweteimgnnglmgvneiqytr 61
 QY 59 QDGANSFYRQDKLFKPTLLADNINPQTGAPGCTDFMTNGVLDVWAMGACNTNNGCY 118
 DB 62 re-naegvgklvt-----takredydgfk 85
 QY 119 RTGAAGDIPPMASARVTFQKYSFTHGKRVVYAKMPVGDMLPAILMPLPEDWYVGMPRS 178
 DB 86 yt-----sarlkqfidskwygkfeakmaipsfrgyvwmfmwsgdntnuyrtpss 135
 QY 179 GEIDIFETIGNRDPFKNTGGEFLGIQKMGSTWMMGPMGMDNRYWLTSLP-----KHSDDN 233
 DB 136 geidfie-----hntune-----kvrgtlnw-----stpdgahahnresn 172
 QY 234 -YGDNEHTEFWMSPGALFEFVYDENQALLDVPYPLIDANPMWVDFEWGKPMPLPOYEND 292
 DB 173 tngldyhllysvemssikvfnng-----gyfev 202
 QY 293 NPMAGCTMLAPEDQNFHILNAVAGGTNGFIPDGCINRGDPALQKPMGNGDMYNDAMK 352
 DB 203 klgvgvngksafinkvfvllmaiggn-----wpgfdvadea--- 239
 QY 353 FFDARGNKKWTWDEGDNNAQVDYIRYTK 382
 DB 240 -fpak-----myldvryvyq 253

RESULT 11

AA67918 standard; Protein; 233 AA.

AA67918;

14-SEP-1995 (first entry)

(1-3)-beta-D-glucan sensitive factor glucanase domain (4-236).

(1-3)-beta-D-glucan sensitive factor glucanase domain (4-236);

antifungal agent; mycosis diagnosis.

Limulus sp.

WO9501432-A.

12-JAN-1995.

29-JUN-1994; 94MO-JP01057.

29-JUN-1993; 93JP-0184403.

(SEK) SEIKAGAKU KOGYO CO LTD.

Iwanaga S, Muta T, Oda T, Seki N;

WPI; 1995-060996/08.

DNA encoding a polypeptide comprising a tetrapeptide motif at least once - which may be used as an antibacterial and antifungal.

Claim 5; Page 40; 51pp; Japanese.

AA67917-R67919 are active fragments of AA67915, a (1-3)-beta-D-glucan sensitive factor, they have a high affinity for the (1-3)-beta-D-glucan found in fungal cell walls. The proteins are therefore useful for clinically diagnosing mycosis, and as antifungal agents for the removal of fungi.

Sequence 233 AA;

Query Match 9.3%; Score 208; DB 16; Length 233;

Best Local Similarity 21.3%; Pred. No. 6.7e-12; Matches 79; Conservative 45; Mismatches 98; Indels 146; Gaps 13;

```

23 QYHIVMODEDFYDGAKEHEV--TATGGSEFOLYTODGANSFVRDGLFIKPTLLAD 80
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 kwqlvnsdeltnglsaswefemngnglmgwnealqyrrf--naqveggklyl----- 52
QY 81 NINPQTGAPGTDFMYNGVLDVWAMYGACTWTDNNGCYRTGACADIPRMSARVTFQKY 140
Db 53 -----takredydgfkyl-----sariktqgdk 75
QY 141 SFTHGKRVYVAKMPVGMILMPALMLEDWYGGWPRSGEIDIIETIGNRPFKNTGSEFL 200
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 76 swkygkkaakmalpsftgyvwmfwmsgdnltyvpsgsgefdle-----hrrlme-- 127
QY 201 GIGKMGSTMHNGPCMDNRYLWLSLP-----KHSDDMN-VGDNFHTFWFDMSPGRLRFV 254
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 128 ---kvggtlhw-----stpdgahhhresntngldyhlysvewmsstkvkfv 172
QY 255 DDEQNALDVPYPLIDANPMWVDFWEMGKFWLPDYENDNPMAGSTNLAPFDQNFHTLNV 314
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 173 ngn-----gyfewkiggyvngkxsfndkfvallm 202
QY 315 AVGGTNGFIDGCTNRGDPALQKPSNGWYNDAMRKFFPARGNWKTWDEGDNNAMQ 374
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 203 a1g9gn-----wpgfdvadea-----fpak-----my 223

```

QY 375 VDYIRVYK 382

Db 224 ldyvrvyq 231

RESULT 12

AA92310 standard; Protein; 341 AA.

AA92310;

01-APR-1999 (first entry)

Streptomyces sp. SGTase protein.

SGTase; beta-1,3-glucanase.

Streptomyces sp.

Key Location/Qualifiers

FT Peptide 1..64

FT Protein /label= signal_peptide

FT Protein 64..341

FT Protein /note= "SGTase"

JP10337184-A.

22-DEC-1998.

09-JUN-1997; 97JP-0151321.

09-JUN-1997; 97JP-0151321.

(DNIN) DAINIPPON INK & CHEM INC.

WPI; 1999-113611/10.

N-PSDB; AAX01742.

New recombinant DNA - used for recombinant production of beta-1,3 glucanase

Disclosure; Page 7-9; 10pp; Japanese.

Sequence 341 AA;

Query Match 7.5%; Score 168.5; DB 20; Length 341;

Best Local Similarity 24.0%; Pred. No. 6.2e-08; Matches 105; Conservative 30; Mismatches 108; Indels 195; Gaps 27;

```

QY 1 MRWTLVVL--CLLRGEGRFT-----DMDQYHIVMODEED-----YFDGA 38
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 37 lrrllvalagalaqag-alcitptlaasvpppsgvtq---viadfdgpkgsqvdgt 92
QY 39 KWQHEVATG-----GENSEFOLYTODGAN-SFVRDGLFIKPTLLADNINPOTGAPF 90
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 93 dwry-atgtygpggsnagtgeltmtanpenvsldgnlrltp----- 136
QY 91 GTDFMYNGVLDVWAMYGACTWTDNNGCYRTG---AAGDIPRMSARVTFQKYSFTHGR 146
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 137 -----trdgagnwtsgrletariddfqpaggltlr----- 165
QY 147 VVYAKMPV---GD---WLPALMML--PEDWYGGWPRSGEIDIIETIGNRDKNTGG 197
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 166 --vearidvprntvtdaakgyrpatfmlqgyrtydywmwpavageldime-----nt-- 213
QY 198 EFLGIGKMGSTMHNG--PGMDNRYLWLSLPKHSDDMNNGDNFHTFWFDMWSENG-----LR 251
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```



```

QY 24 YHIVWDEDFEDYGA-----KMOHEVTATGAGNSEFQLYTDGANSFVRDGLFIKPTLL 78
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 fvlvwtldttagngtspnqnnw-niltgnlnvaeegetysstaavqlsggs-----tl- 80
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 ADNINPOTGAPCTGPMYNGVLDVWAMAGACTNTDNNGCYRTGAADIPPMASARVTFQ 138
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 81 -qlvwrtdsskgt-----stfgswt-----sgrles-- 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 139 KTSFT--HRRVY-VHAKMPVG-----DMLMPAIMMLPEDWY--GMPRSEIDITIER 186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 KYCfFpaagkvtrleaaifgsnaganakkgilwpaftwmlgslqypgswpcgeidmet 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 IGRNDFKNKGGEFL-----GIQKMGSTWMPGPR-----GMDNRMYLTLSPKSDPMNYG 235
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 v---dgatqghctlhcdvypgslcneqngf-ggprvlnanvdhavrveldrtpssq-s 220
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 DNFHFWFMDSPNGLEPFYVDENQALLDVYPYLIDANPMWVDFWEMGKFWLPQYENDNPM 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 en-----lwsldgtlyfqltgs-----rignqyv 246
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      296 AGSTNLAPRDQNFHFLNVAVGSTNGFIPDGCINRGDPAALQKPMNSGDMYNDAMRKF 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      247 ---mla--hsplflllnvaavg-----nwpqnp-----nsacl---- 275
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 ARGNNKWTWDEGDNNAMQVDYIRVY 381
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 -----dgysmmevyvaqy 290
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

AAW77311
ID AAW77311 standard; Protein; 424 AA.

AAW77311;

10-DEC-1998 (first entry)

Phaffia rhodozyma endo-1,3(4)-beta-glucanase.

Phaffia rhodozyma endo-1,3(4)-beta-glucanase; beta-glucan; brewing;
wine-making; protoplast; pigment; colourant; flavourant.

Phaffia rhodozyma.

WO9836056-A1.

20-AUG-1998.

13-FEB-1998; 98WO-DK00057.

14-FEB-1997; 97DK-0000161.

(NOVO) NOVO-NORDISK AS.

Bang M, Sandal T;

WPI: 1998-506310/43.

N-PSDB; AAV59143.

Enzyme with endo-1,3(4)-beta-glucanase activity - which is useful
in, e.g. brewing, wine-making, preparation of protoplasts from
yeasts, or preparation of pigments, colourants or flavourants

Claim 2; Pages 27-28; 36pp; English.

The present sequence represents the filamentous fungus *Phaffia*
rhodozyma (CBS No. 6938) endo-1,3(4)-beta-glucanase sequence.
The enzyme may be used for modification or degradation of material
which contains beta-glucan. Therefore the enzyme is claimed to be
useful in, e.g. brewing, wine-making, preparation of protoplasts
from yeasts, or preparation of pigments, colourants or flavourants.
The enzyme is also claimed to be useful in the preparation of
pharmaceuticals.

```

XX
SQ Sequence 424 AA;
Query Match 5.4%; Score 121.5; DB 19; Length 424;
Best Local Similarity 21.5%; Pred. No. 0.0024;
Matches 59; Conservative 25; Mismatches 98; Indels 93; Caps 14;
QY 108 ACTNTDNNGCYRTGAADIPPMASARVTFQKYSFTIGRV---VHAKMPVGDMLMPAIV 164
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 alvngkynavlavdtlgnvqgrka-vrlhsyllnggillladlvh-mpbgcglwpaaw 216
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 MLPEDWYVGMPRSGEIDITETIGNRDFKN-----TGGEFFGIQ 203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 snpbd-----wpnkgeidllegthswdrngsvhtsdgcltprngasavlttgsfvtn 271
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 KMG-STWHPGPMDDNRXYLTLSPKHSDDNNYGDNFH---TFWDMSPNGLEPFYVDE 257
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 272 casyatangcgqres-----ashqaygepfngnggyvymkwdtagisvffpr 321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 NQALLD---VYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTNLAPRD--QNFHFI 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 322 naipadltgvpilp-----etwgcpm-----gnfp---stscptkfkfkdhtli 362
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 LNVAVGSTNGFIPDGCINRGDPAALQKPMNSGDMY 346
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 intfcg-----dwansdww 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: October 11, 2001, 15:51:39
Job time: 202 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:52:04 : Search time 19.14 Seconds
(without alignments)
413.098 Million cell updates/sec

Title: US-09-596-101a-3

Perfect score: 2240
Sequence: 1 MKMTLVVLCILFEGEFAFTD.....DDEGNNAMQVYIRYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	15.9	321	2	US-08-712-072C-3 Sequence 3, Appl
2	351.5	15.7	262	1	US-08-392-828C-37 Sequence 37, Appl
3	351.5	15.7	262	3	US-09-330-945-37 Sequence 37, Appl
4	344.5	15.4	276	2	US-08-712-072C-4 Sequence 4, Appl
5	307	13.7	306	2	US-08-824-707-2 Sequence 2, Appl
6	286.5	12.8	285	2	US-08-712-072C-5 Sequence 5, Appl
7	272.5	12.2	422	2	US-08-712-072C-2 Sequence 2, Appl
8	208	9.3	654	1	US-08-392-828C-2 Sequence 2, Appl
9	208	9.3	654	3	US-09-330-945-2 Sequence 2, Appl
10	129.5	5.8	292	2	US-08-737-526-4 Sequence 2, Appl
11	129.5	5.8	292	4	US-09-098-580-4 Sequence 4, Appl
12	112.5	5.0	738	1	US-07-985-458-3 Sequence 3, Appl
13	110	4.9	243	3	US-09-286-690-10 Sequence 10, Appl
14	106	4.7	238	1	US-09-286-690-7 Sequence 7, Appl
15	100	4.5	237	1	US-08-103-998-4 Sequence 4, Appl
16	99	4.4	279	3	US-09-286-680-9 Sequence 9, Appl
17	98.5	4.4	239	1	US-08-103-998-2 Sequence 2, Appl
18	92.5	4.1	312	4	US-09-216-295-21 Sequence 21, Appl
19	92.5	4.1	371	4	US-09-104-308-1 Sequence 1, Appl
20	90.5	4.0	242	3	US-09-286-690-11 Sequence 11, Appl
21	90.5	4.0	429	1	US-08-745-977-4 Sequence 4, Appl
22	90.5	4.0	429	3	US-09-040-699A-4 Sequence 4, Appl
23	90.5	4.0	552	3	US-09-120-365-5 Sequence 5, Appl
24	90.5	4.0	552	4	US-09-515-039-5 Sequence 5, Appl
25	89.5	4.0	491	1	US-08-206-176-4 Sequence 4, Appl
26	89.5	4.0	553	2	US-08-661-052-16 Sequence 16, Appl
27	89	4.0	750	6	5457037-3 Patent No. 5457037

28	89	4.0	751	6	5457037-5
29	89	4.0	822	3	US-08-941-445A-17
30	87.5	3.9	1052	4	US-09-255-502-7
31	87	3.9	635	2	US-08-797-366-5
32	87	3.9	635	2	US-08-956-268-5
33	87	3.9	854	2	US-08-928-692-17
34	86	3.8	534	4	US-09-124-541-1
35	85.5	3.8	242	3	US-08-814-823-6
36	85	3.8	215	1	US-08-451-409A-1
37	84	3.8	485	4	US-08-749-391-2
38	84	3.8	485	4	US-09-390-200-2
39	84	3.8	560	3	US-08-814-823-6
40	84	3.8	560	3	US-08-814-823-6
41	84	3.8	799	3	US-08-941-445A-15
42	83.5	3.7	289	2	US-08-580-545B-4
43	83.5	3.7	289	4	US-09-262-653A-4
44	83.5	3.7	289	4	US-08-867-484A-2
45	83.5	3.7	313	3	US-08-508-761B-34

ALIGNMENTS

RESULT 1
US-08-712-072C-3
Sequence 3, Application US/08/12072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: el3b, Bacillus circulans
US-08-712-072C-3

Patent No. 5457037
Sequence 17, Appl
Sequence 7, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 1, Appl
Sequence 8, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 15, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 34, Appl

```

Query Match 15.9%; Score 357; DB 2; Length 321;
Best Local Similarity 26.2%; Pred. No. 2.8e-26;
Matches 100; Conservative 46; Mismatches 86; Indels 150; Gaps 13;

Oy 25 HIVWDEF-----DYFDKAKHREVTATGGNSEPOLYTODGANSFYRDKLFIK 74
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 64 NLIWDEFNETTLDTSKNNTFTGYLLNDPRAWGNGNAELQHTNSTONVYQDRLNIK 123
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 75 PYLLADNINPOTGAFGFDPMYNGVLDVWAMYGACTNTDNNCYRTGAAGDIPPMASRV 134
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 124 --AMNDSKSPQOP-----NRXAQS-----SGKI 145
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 135 RRFQKSTFHGNYVYHAKKPVGDMLMPALWMLPEDMVYGGPRSGEIDIIETIGNRDEFN 194
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 146 NKDKLTLSTKYGVDRPRAKPTGDIQWVPMPLMKDSYVGTMAASEIDMEARGLPGSV 205
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 195 TGGERTLGIOKMGSTHWGSGMDNRYTWLTLSLKHSDDNN-----YGDHFHFWDMSN 248
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 206 SG-----THFGQMPVNO-----SSGDYHFPFGQTFANDYHIVSYWEED 247
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 249 GLRPFVDE-----NOALLDVPYPLIDANPMWVDFWEMGKFWLPOYENDNFWAGTNTLA 302
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 248 NIKWVVDGKFFKYVINO-----QWYSNAANNPV-----A 277
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 303 PPDQNFHILNAVAGT--NGEIPDGCINRGDPLQKFWMSGDWYNDAMKFFPARGNW 360
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 278 PFDEFYFLIMNLVAGNFGDGRTP---NASDIPA----- 308
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 361 KWTWDEGDNNAMQVDYIRVYK 382
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 309 -----TMQVDYVRVYK 319
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
US-08-392-828C-37
: Sequence 37, Application US/08392828C
: Patent No. 5795962
: GENERAL INFORMATION:
: APPLICANT: IWANAGA, SADAOKI
: APPLICANT: MUTA, TATSUSHI
: APPLICANT: SEKI, NORIKI
: APPLICANT: ODA, TOSHIO
: TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
: STREET: 53 STATE STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/392,828C
: FILING DATE: 28-FEB-1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: CAMPBELL, PAULA A
: REGISTRATION NUMBER: 32,503
: REFERENCE/DOCKET NUMBER: FJN-033
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 37:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 262 amino acids
: TYPE: amino acid

```

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..262
: OTHER INFORMATION: /note= "BGI A1 SEQUENCE (FIGURE 2)"
: US-08-392-B28C-37

Query Match 15.7%; Score 351.5; DB 1; Length 262;
Best Local Similarity 26.2%; Pred. No. 7.6e-28;
Matches 100; Conservative 44; Mismatches 89; Indels 149; Gaps 14;

QY 25 HIVWDEDF--YFDGAKQHEV-----TATGGNSEFQLYTODGANSFVRGKLFK 74
:::||||: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 4 NLIWDEDFNGTLLDTSKKNYETGGYLLNNDPATFWGNAELQHTYNSIQNVYDGSKLNK 63
:::||||: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 75 PTLADNINPOTGAFGFGDEMNGVLDVWAMYGACTNTDNNGCRTGAAGDIPPMASRV 134
: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 64 A-----NDKSKFPQDP-----NRYAQS-----SGKI 86
: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 135 RFEQKYSTHGYYVVAHAMPYGDMLPAILMLPEDWYGGWPRSGEIDIIETGNRDPKN 194
:::||||: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 87 NKDLSLSTKYGVDFRALKPTGDSWPAALMLPKDSYIGTVAASGSEIDVMEGRKRLPSV 146
:::||||: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 195 TGEERLGIQKMGSTWHPGMDNRRYWLTSLPKHSDDN-----YGDNFTFWFDSPN 248
: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 147 SG-----TIFSGQMPVNO-----SSGGDIHPPEGQTFANDYHYVSWEED 188
: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 249 GIREFYVDE-----NQALLDVPYPLIDANFMMWVDFWEGKFWLPQYENDNFWAGGTILA 302
: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 189 NIKWVYDKEFEYKYVNO-----QWYSPAAPPNNPN-----A 218
: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 303 PPDQWFHFLINVAVGT--NGFIPDGCINRGDPALQKPMWSNGDWYNDAKMKFFDARGNW 360
|||: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 219 PFDEFYFLIMNLAVGANFDGGRTP---NASDIPA-----A 249
: |::| | |::| | |::| | |::| | |::| | |::| | |
QY 361 KWTWDEGDNNMAQVDYIRYK 382
|||: |::| | |::| | |::| | |::| | |::| | |::| | |
DB 250 -----TMQVDYIRYK 260)
: |::| | |::| | |::| | |::| | |::| | |::| | |

RESULT 3
US-09-330-945-37
: Sequence 37, Application US/09330945
: Patent No. 6077946
: GENERAL INFORMATION:
: APPLICANT: IWANAGA, SADAOKI
: APPLICANT: MUTA, TATSUSHI
: APPLICANT: SEKI, NORIAKI
: APPLICANT: ODA, TOSHIO
: TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
: TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
: ADDRESSSEE: THIBEAULT, LLP
: STREET: 125 HIGH STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/330,945
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:

```


APPLICANT: Hedegaard, Lisbeth
APPLICANT: Haikler, Torben
APPLICANT: Aasenjo, Juan
APPLICANT: Sava, Demetris
TITLE OF INVENTION: No. 591688el enzyme with beta-1,3-glucanase activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 591688el No. 591688elisk of No. 591688th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,707
FILING DATE: 14-April-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4290, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-824-707-2

Query Match 13.7%; Score 307; DB 2; Length 306;
Best Local Similarity 26.2%; Pred. No. 3.1e-23;
Matches 95; Conservative 36; Mismatches 97; Indels 134; Gaps 13;

QY 26 IWNODEFDYDGA-----WQHEVTATGGGSEFQLYTQDANSFYRQKLFIRPTLLAD 80
DB 68 IANSDEFGAAGSAPNDVWNNHETGAGGMAELQNTTSSVNSAL-DGQ----- 116
QY 81 NINPOTGAPFTDEMNGVLDVWMAVYGACTDNNNGCYRTGAGDIPRMSARVTFQKY 140
DB 117 -----GNLVITL-----QESDGSY-----TSARLTQGNV 142
141 SFTHGRVVAHAKMPVGMILPAILMLPBDWVYGGMPRSGEIDITETGNRDKFTGSEFL 200
DB 143 QPOFGRIEARLQIPRGOGISAFMMVGNALPDTPPTSGEIDIMENVGNAAHEVHG----- 198
QY 201 GLOKMGSTMHGPGMD-DNRWLTSLPKHSDWYGNDFHTFPMFWSNGIRREFVDDENO 259
DB 199 -----TVH-GPRTSGN--GIMGTYOHFGQMSFADDFHFGIDWTFGEITWLD--GQ 246
QY 260 ALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTNLAPDQNFHFLINAVAGGT 319
DB 247 EYHVTYADVAGANQ-----VFDDQFFFLINVAIG- 277
QY 320 NGFIPDCCINRGDPALQKPSNSGDWYNDAMRKFFDARGNKKWTWDEGDNNAQVDYIR 379
DB 278 -----QMPGNPDATTFPP-----QQMKVDYVR 299

RESULT 6
US-08-712-072C-5

Sequence 5, Application US/08712072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
AMBI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium keratolyticus
US-08-712-072C-5

Query Match 12.8%; Score 286.5; DB 2; Length 285;
Best Local Similarity 27.0%; Pred. No. 3.4e-21;
Matches 103; Conservative 32; Mismatches 100; Indels 147; Gaps 15;

QY 10 LIFGSGFATFMDQYHAYWODEFDY---FDGAKWQHEVTATGGN---SEFQLY-TODGA 62
DB 8 LLNATVATFD---YELMSDEFNSSGGFDSKWKYADRGTAAKKYKMSLPAVYSQDS 64
QY 63 NSFVDEGLFKFTPLLDNINPQTCAPFTDPMYNGVLDVWMAVYGACTDNNNGCYRTGA 122
DB 65 NLVLR-----MDNAVY----- 75
QY 123 AGDIPRMSARVTFQKYSFTHGRVVAHAKMPVGMILPAILMLPBD-WVYGGMPRSGEI 181
DB 76 AGDPYAVHAGVSKSGKTSMTYGVKVEYAKFTQGRGSPWALWMPDPATAYGWSGCGEI 135
QY 182 DIIFETGRNDEKNTGEEFLGIQKMGSTMHGPGMDNRYWLTSLPKHSDWYGNDFHTF 241
DB 136 DSMEHVNNESV-----MYHTIHNGSVYTNANG--GSTASKSKATVYTFD-YNYX 179
QY 242 WEDMSPNGLREFVDDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTNL 301
DB 180 TWMSPNDRIFRY---NNSL-----QYTVARVSGGGTQ 210
QY 302 APFDQNFHFLINAVG-GTNGFIPDCCINRGDPALQKPSNSGDWYNDAMRKFFDARGNW 360

Db 211 WPEFVPEYLLINQAGAGWPGATINADL-----PFS----- 241
 QY 361 KMTWDEGDNNAQVYRYRK 382
 Db 242 -----MOVDYRYRK 251

RESULT 7

US-08-712-072C-2
 ; Sequence 2, Application US/08712072C
 ; Patent No. 5925541

; GENERAL INFORMATION:
 ; APPLICANT: Jack Goldstein, Alex Zhu and Ian Leng
 ; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
 ; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amster, Rothstein & Edenstein
 ; STREET: 90 Park Avenue
 ; CITY: New York
 ; STATE: NY

; COUNTRY: US
 ; ZIP: 10016

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/712,072C

; FILING DATE: 11-SEP-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Bogosian, Elizabeth A.

; REGISTRATION NUMBER: 39,911

; REFERENCE/DOCKET NUMBER: 63475/97

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 286-0854 or 286-0082

; TELEFAX: TWX 710-581-4766

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 422 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHEICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

US-08-712-072C-2

Query Match 12.2%; Score 272.5; DB 2; Length 422;
 Best Local Similarity 27.2%; Pred. No. 1.6e-19;
 Matches 105; Conservative 29; Mismatches 97; Indels 155; Gaps 17;

QY 10 LLFEGFAFTMDQYHIWODEFDY---FDGAKQHEVTATGSGSEPOLYODGANSFV 66
 Db 45 LLNATVATFTD---YELTWSDEFNSSGGFDSTKMSYADRGTVAMNK--YMTSDGSSNLVL 99
 QY 67 RDGLFIKPTLLADNINPOTGAPFCTDFMYNGVLVWVWAMYGACTTNDNGCYRTGAAGDI 126
 Db 100 R-----MDNAV-----AGD- 108
 QY 127 PPMSARRTYOKYSFTIGRVVYNAK-----MFV---GDMLPAIWMLEPD-WYTGGMFR 177
 Db 109 FVAHAGGVCKMSKESMTYKVEVRAKFTGVSLPAYAGRGSPAIIMMPEPATAYAGWPS 168

QY 178 SGEIDIIEITIGNRDKTNGGEFLGIQKMGSTMHGPGWMDNRYWLTSLPKHSDMANYGDN 237
 Db 169 CGEIDSMENHNESV-----MYHTLHNSVTNANG---GSTASKSATYTTD- 212
 QY 238 FHTFWDSPNGLRFFVDDENQALDVPYPLIDANFMYVDFEMGKRPWLPQTEENDNPWAG 297
 Db 213 YNLTYMTWSPNDIRFYV---NNSL-----QYTYARVSG 243

QY 298 GTNLAPFQONFETLNVAVG-GTNGFIPDGCINRGDPALOKPWSNGDYNAMRKFFDA 356
 Db 244 GTQMPFDPYFLLINQAGAGWPGATINADL-----PFS----- 278

QY 357 RGNMKTWDEGDNNAQVYRYRK 382
 Db 279 -----MOVDYRYRK 288

RESULT 8

US-08-392-828C-2
 ; Sequence 2, Application US/08392828C
 ; Patent No. 5795962

; GENERAL INFORMATION:

; APPLICANT: IWANAGA, SADAOKI

; APPLICANT: MUTA, TATSUSHI

; APPLICANT: SEKI, NORIARI

; APPLICANT: ODA, TOSHIO

; TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &

; STREET: 53 STATE STREET

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/392,828C

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: CAMPBELL, PAULA A

; REGISTRATION NUMBER: 32,503

; REFERENCE/DOCKET NUMBER: FJN-033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 654 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-392-828C-2

Query Match 9.3%; Score 208; DB 1; Length 654;
 Best Local Similarity 21.5%; Pred. No. 1e-12;
 Matches 79; Conservative 45; Mismatches 98; Indels 146; Gaps 13;

QY 23 OYHIWODEFDYFDGAKQHEV--TATGGNSEFOLYTOGANSFVRDGKIFKPTLLAD 80
 Db 5 KQOLWSDFTNGISSDMEFEKNGNGNNEIQYRRE--NAOVEGGKLV----- 55
 QY 81 NINPOTGAPFGTDFMYNGVLVWVWAMYGACTTNDNGCYRTGAAGDIPPAMSAVRRTFOKY 140
 Db 56 -----TAKREDYDGFKYT-----SARLKTQFOFK 78

Query Match 5.8%; Score 129.5; DB 2; Length 292;
 Best Local Similarity 22.0%; Pred. No. 3.1e-05;
 Matches 85; Conservative 40; Mismatches 110; Indels 151; Gaps 22;

24 YHIVODEFDYDGA-----KMOHEVTATGGNSEFQLYTODGANSFYRDKLFKIPILL 78
 28 FNLVWTDTPAANGTSPNQNNW-NITGNLVNVAEQETYSSTANVOLSGS-----TL- 80
 79 ADNINPOTGAPGTDFMYNGVLVWAMYGACTNTDNNGCYRTGAGDIPPMASARVTFQ 138
 81 -QLVPMRDSKGT-----STFGWT-----SGRLSS-- 105
 139 KYSET--HGRVY-VHAKMPV-----DWLMPAIVMLPEDVY--GGMPRSGEIDILET 186
 106 KYTFTPAAGKYTRLEAIRFGSNAQANKQIWPAPFMTGLDRLPGSGSWPNCGEIDIMET 165
 187 IGRNDEKNTGGEFL-----GIQKMGSTAMGP-----GMDNRYWLSLPKHSDDNNY 235
 166 V--DGAATGHTLHCDYTPGICNEGNGI--GGPVNIANVNDHMAVREIDTPSSW--- 218
 236 DNFHTFMDSPNGRLRFVDDENQALLDVPYPLIDANPMWVDFEWGKFWLPQYENDNPM 295
 219 ---QSETLWMSLDGTIYQITGS-----RIGNQGVW 246
 296 AGGTNLAFPDQNFHILNAVVGCTNGFIIPDCINRGDPALQKPSNGDWTNDAMRKFFD 355
 247 ---NNIA--HSPLEFFILNAVVG-----NMPGNP-----NSATL----- 275
 356 ARGNWKWTWDEGDNNAQVYIRY 381
 276 -----DGYGSMMEVGYVQY 290

RESULT 11

US-09-580-4
 ; Sequence 4, Application US/09098580

Patent No. 6140096

GENERAL INFORMATION:

APPLICANT: Kofod, Lene Venke

APPLICANT: Andersen, Lene No. 6140096doe

APPLICANT: Kauppinen, Markus Sakari

APPLICANT: Christgau, Stephan

TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase

TITLE OF INVENTION: Activity

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6140096 No. 6140096disk of No. 6140096th America, Inc.

STREET: 405 Lexington Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/098,580

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/737,536

FILING DATE: 08-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Valeta, Gregg A

REGISTRATION NUMBER: 35,127

REFERENCE/DOCKET NUMBER: 4174,204-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

TELEX:
 INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 292 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-580-4

Query Match 5.8%; Score 129.5; DB 4; Length 292;
 Best Local Similarity 22.0%; Pred. No. 3.1e-05;
 Matches 85; Conservative 40; Mismatches 110; Indels 151; Gaps 22;

24 YHIVODEFDYDGA-----KMOHEVTATGGNSEFQLYTODGANSFYRDKLFKIPILL 78
 28 FNLVWTDTPAANGTSPNQNNW-NITGNLVNVAEQETYSSTANVOLSGS-----TL- 80
 79 ADNINPOTGAPGTDFMYNGVLVWAMYGACTNTDNNGCYRTGAGDIPPMASARVTFQ 138
 81 -QLVPMRDSKGT-----STFGWT-----SGRLSS-- 105
 139 KYSET--HGRVY-VHAKMPV-----DWLMPAIVMLPEDVY--GGMPRSGEIDILET 186
 106 KYTFTPAAGKYTRLEAIRFGSNAQANKQIWPAPFMTGLDRLPGSGSWPNCGEIDIMET 165
 187 IGRNDEKNTGGEFL-----GIQKMGSTAMGP-----GMDNRYWLSLPKHSDDNNY 235
 166 V--DGAATGHTLHCDYTPGICNEGNGI--GGPVNIANVNDHMAVREIDTPSSW--- 218
 236 DNFHTFMDSPNGRLRFVDDENQALLDVPYPLIDANPMWVDFEWGKFWLPQYENDNPM 295
 219 ---QSETLWMSLDGTIYQITGS-----RIGNQGVW 246
 296 AGGTNLAFPDQNFHILNAVVGCTNGFIIPDCINRGDPALQKPSNGDWTNDAMRKFFD 355
 247 ---NNIA--HSPLEFFILNAVVG-----NMPGNP-----NSATL----- 275
 356 ARGNWKWTWDEGDNNAQVYIRY 381
 276 -----DGYGSMMEVGYVQY 290

RESULT 12

US-07-985-458-3
 ; Sequence 3, Application US/07985458

Patent No. 5344777

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;

APPLICANT: Takemura, Hiroshi;


```

      1  FILING DATE: 19921203
      2  CLASSIFICATION: 435
      3  PRIOR APPLICATION DATA:
      4  APPLICATION NUMBER: US 07/658,221
      5  FILING DATE: 20-FEB-1991
      6  APPLICATION NUMBER: 73440/1990
      7  FILING DATE: 26-MAR-1990
      8  ATTORNEY/AGENT INFORMATION:
      9  NAME: Goodman, Herbert
     10  REGISTRATION NUMBER: 17081
     11  REFERENCE/DOCKET NUMBER: 910134/MG
     12  TELECOMMUNICATION INFORMATION:
     13  TELEPHONE: (212)972-1400
     14  TELEFAX: (212)370-1622
     15  TELEX: 236268
     16  INFORMATION FOR SEQ ID NO: 3:
     17  SEQUENCE CHARACTERISTICS:
     18  LENGTH: 738 amino acids
     19  TYPE: amino acid
     20  TOPOLOGY: unknown
     21  MOLECULE TYPE: protein
     22  FEATURE:
     23  NAME/KEY: MATURE PEPTIDE
     24  LOCATION: 36 to 738
     25  IDENTIFICATION METHOD: N-terminal sequences of the
     26  IDENTIFICATION METHOD: purified protein having a molecular weight of about
     27  IDENTIFICATION METHOD: 72,000
     28  ORIGINAL SOURCE:
     29  ORGANISM: Acetobacter altoacetiigenes
     30  STRAIN: MH-24
     31  PUBLICATION INFORMATION:
     32  AUTHORS: Tamaki, Toshimi;
     33  AUTHORS: Fukaya, Masahito;
     34  AUTHORS: Takemura, Hiroshi;
     35  AUTHORS: Tayama, Kenji;
     36  AUTHORS: Okumura, Hajime;
     37  AUTHORS: Kawamura, Yoshiya;
     38  AUTHORS: Nishiyama, Makoto;
     39  AUTHORS: Horinouchi, Sueharu and
     40  TITLES: Cloning and sequencing of the gene cluster
     41  TITLES: Encoding two subunits of Membrane-Bound
     42  TITLES: Alcohol Dehydrogenase from Acetobacter
     43  JOURNAL: Biochimica et Biophysica Acta.
     44  VOLUME: 1088
     45  PAGES: 292-300
     46  DATE: 1991
     47  524-07-985-458-3

```

```

Db 428 LGGHNFAMASPP-----KTGLVYIPAAQVPEFLYNQVGGFFPHDSSNNLGLDMKKVGLP 482
QY 245 WSPNGREFVEVDENQALLDVPYPLIDANPMWVDEW---WGRPWLPOYENDPMAGTNL 301
Db 483 DSPEAKQAFVKDLK-----GRTVAMDPOKQAEAW--RVDHKGPMMNGJLA 525
QY 302 APFDQNFHFIILN---VANGTNG 321
Db 526 TGGDLLFQGLANGEFHAYDATING 548

RESULT 13
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; NUMBER OF SEQ. ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-286-690-10

Query Match 4.9%; Score 110; DB 3; Length 243;
Best Local Similarity 21.9%; Pred. No. 0.0023;
Matches 60; Conservative 42; Mismatches 88; Indels 84; Gaps 18;

QY 71 LEIKRTLLADNINPOTGAPFGTDF-MYNGVLDDVMAMYGACNTDNNNGCYRFGAAGDIPPA 129
Db 16 LFLSLSTRASASADTGGSGFYEPFNNTY--TGLWOKADGYSNGNPNFNC--TWKANNVSKT 71
QY 130 MSARYR-----TEQKY-----SFTGRVVYHATKPYGDWMLMPA--IWMLEPEDVY 172
Db 72 SLGEMRLSLTSPSYNKKFPGCGENRSVQTYGYGLYEVMK-----PAKNQIVSSFFTY 123
QY 173 GGMPSRG---EIDIIETIGNRDEKNTGGEFLGLOKMSYIMHMGWDNRKYWLTSLPKH 228
Db 124 TG-PTDGTPEWDEIDI-----EFLG--KDTTKVOF-----NYITNGVGNH 159
QY 229 SDDMWY---DNFHTWFMDSPNGLREFVDE--NQALLDVPY--LIDANPW---WYD 277
Db 160 EKIYVLMGDAANSYTYFAFDWQPNSTKMYVDGQLKHTATYTOIPOTPGKIMMLNMGAGVD 219
QY 278 FWEWGRPWLPOYENDPMAGSTNLAPFDQNFHFI 311
Db 220 -----EWLGSY-----NGYIPLSRSLHW 238

RESULT 14
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05

```

EARLIER APPLICATION NUMBER: US 60/027,882
EARLIER FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: PCT/US97/17811
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 238
TYPE: PRT
ORGANISM: Bacillus polymyxa
US-09-286-690-7

Query Match 4.7%; Score 106; DB 3; Length 238;
Best Local Similarity 22.0%; Pred. No. 0.0056;
Matches 62; Conservative 32; Mismatches 92; Indels 96; Gaps 16;

QY 25 HIWODEFDYFDGAKMOHEVATGGSSEFQLYTODGAN-SFVRDQKLFIRPTLLADNIN 83
28 NVTW-EPLSYFNSSTWQ---KADQSYNGQMPNCTWRANNVFTNDGKIKLSLTPANN-- 81
DB 84 PQTGAPGTDFMTNGVLDVWAMYGACTNTDNGCYRTGAAGDIPRMSARVTFQKYSF- 142
82 -----KFDG---GEYRSTNNGY-YGLYEYSMPKPAKNTGIYSFETYTGP 121
QY 143 THGVVYVHAKMPVQDWLMPRAIMLPEDWYVGGWPRSGEIDIEITGNRDFKNTGGEFLGI 202
122 SHG-----TQM-----DEIDI-----EFLG- 136
QY 203 OKMGSTWNGGMDNRXYWLTSLPKHSDDNNGY---DNFTFWEDMSPNGLRPFVD-- 256
137 -KDTKQV-----NYTNGVGGHEKIIIMLGSDASTSFITVAFDMPQGIKTVYDVL 188
DB 257 ENQALDVPYPLIDANPWVDFWENG---KFWLPQYENDNP 294
189 KHRATYNIPT---STPGKIMMNLNMGTVDSWLGSYNGANP 225

RESULT 15
US-08-103-998-4
Sequence 4, Application US/08103998
Patent No. 5470725
GENERAL INFORMATION:
APPLICANT: Bottiss, Rainer
APPLICANT: Hofemeister, Jurgen
APPLICANT: Thomsen, Karl Kristian
APPLICANT: Olsen, Ole
APPLICANT: Von Wettstein, Dietrich
TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 1800 Diagonal Road, P.O. Box 299
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,998
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/773,652
FILING DATE: 16-Oct-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30307/123

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-103-998-4

Query Match 4.5%; Score 100; DB 1; Length 237;
Best Local Similarity 20.6%; Pred. No. 0.023;
Matches 42; Conservative 21; Mismatches 71; Indels 70; Gaps 9;

QY 156 GDMW-PAIMWLPEDWYV-GWPRSGEIDIEITGNRDFKNTGGEFLGIQKMG----- 206
26 GSVTFEPLSTYNPSTWEKADQSYNGGVNCTWRANNVFTNDGKIKLSLTPANN-- 85
DB 207 ---STMWGPQMD-----NRYWLTSLPKHSDDW-----NY 234
86 EYRSTNIYGYGLYEYSMPKPAKNTGIYSFETYTGPAGHTQWDEIDIEFLGKDTKQVFN 145
QY 235 GDN-----FHTFWEDMSPNGLRPFVD--NQALDVPYPLIDANPW 274
146 YTNAGNHEKFPADLGFDAANAYHTYAFDQPNISIKWYDGLKHTATQIP---AAG 200
QY 275 WVDWENG---KFWLPQYENDNP 294
201 KIMMNLNMGTVDSWLGSYNGVNP 224

Search completed: October 11, 2001, 15:52:05
Job time: 208 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:55:27 ; Search time 25.22 Seconds

(without alignments)
1159.835 Million cell updates/sec

Title: US-09-596-101A-3

Sequence: 1 MRRTLVVLCILFEGGFATD.....DDEGDNNAMQVDYIRVYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Maximum number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 68:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	31.8	499	2 JC6141	beta 1,3-glucanase
2	369.5	16.5	877	2 JN0772	glucan endo-1,3-beta
3	353	15.8	1324	2 T18265	endo-1,3(4)-beta-g
4	351.5	15.7	682	2 JQ0420	beta-1,3-glucanase
5	344.5	15.4	286	2 S48201	licheninase (EC 3.
6	314	14.0	467	2 JC6150	gram-negative bact
7	284	12.7	642	2 B72428	lamellarinase - The
8	209	9.3	673	2 A49878	coagulation factor
9	179.5	8.0	466	2 T35164	probable secreted
10	148	6.6	364	2 T17584	probable beta-1,3-
11	138.5	6.2	294	2 D70525	probable beta-1-
12	132	5.9	419	2 T39920	probable glucanase
13	125.5	5.6	1144	2 A75132	hypothetical prote
14	125	5.6	857	2 B69798	hypothetical prote
15	116.5	5.2	878	2 A83748	endo-beta-N-acetyl
16	115	5.1	284	2 T06270	xyloglucan endo-1,
17	112.5	5.0	738	2 S14270	alcohol dehydrogen
18	112.5	5.0	954	1 S20907	endo-1,4-beta-xyla
19	111	5.0	282	2 A85354	hypothetical prote
20	110.5	4.9	856	2 T00345	avicelase III - As
21	110	4.9	243	1 S15388	licheninase (EC 3.
22	109	4.9	781	2 A43866	neuraminidase - Vi
23	109	4.9	807	2 B82158	neuraminidase VC17
24	108	4.8	282	2 T02354	xyloglucan endo-1,
25	108	4.8	636	2 T37843	probable beta-gluc
26	108	4.8	720	2 S61143	KRE6 protein - Yea
27	107.5	4.8	758	2 T48815	mixed-linked gluc
28	107.5	4.8	772	2 T02098	probable phosphati
29	106	4.7	238	1 S19012	licheninase (EC 3.

30	106	4.7	742	2 A49340	alcohol dehydrogen
31	106	4.7	782	1 G64157	probable organic s
32	105	4.7	802	2 A36910	xylinase, beta(1,3
33	104.5	4.7	239	1 A29091	licheninase (EC 3.
34	103.5	4.6	851	2 H84053	endo-beta-1,3-1,4
35	103	4.6	289	2 F71402	xyloglucan endo-1,
36	102.5	4.6	276	2 T40453	licheninase (EC 3.
37	102.5	4.6	334	1 S23498	licheninase (EC 3.
38	101.5	4.5	984	2 T44496	cellulose 1,4-beta
39	100.5	4.5	277	2 S71222	xyloglucan endo-1,
40	100.5	4.5	292	2 T06201	xyloglucan endo-1,
41	99	4.4	564	2 T49073	hypothetical prote
42	99	4.4	1087	2 S41797	cellulose 1,4-beta
43	98	4.4	1048	2 C86189	protein T25N20.11
44	98	4.4	1196	2 A29130	beta-amyase (EC 3
45	97.5	4.4	611	2 S60040	alpha-amyase (EC

ALIGNMENTS

Query Match	Score	DB 2:	Length	499;
Best Local Similarity	42.78;	Pred. No.	4.7e-50;	
Matches 158;	Conservative	48;	Mismatches	88;
			Indels	76;
			Gaps	15;
QY	26	IWMODEFDYEDGAKMOHEVTAATGSGSEFQLYTODGANSFYVADGKLFKPTLLADNINQ	85	
Db	179	LIFPEEDSFMDLMEHEMTATGGGMEFEYTTNNRSYSYVADGKLFKPTLTQKL---	235	
QY	86	TGAFFGIDPMYNGYLDVYAMTGA--CTYTDNNGCYRTGAAG-TPPAMSARRTQKXSF	142	
Db	236	-----GECSSLSGTFDLMGSSPANLCTGNANYGCSRTSSNNLMLPDISARITYESPSF	290	
QY	143	THGRVYVHAKKPVGDWMLPALWMLPEDWVYGMPRSGEIDIIETIGNRDFKRTGEGFGLI	202	
Db	291	KYGLVEAKLPTGDMPLPALWLPKHNGYGEWPSGELDVESGNADIKDADLSAGV	350	
QY	203	QKMSSTMHMGWDNDRYMLSLPKSHSDMNYGDNFRTFWFDMSPNGLRFVVDENQALL	262	
Db	351	DQMSSTMHMGFWPLNGV-----PK-----THATKFYVDDE-----	381	
QY	263	DVPPPLIDANRWWDPMWEGKPYLPQYEN-----DNPAGGYN-LAPPDQNHFLINAV	316	
Db	382	---LLNWDV-ATGFMDLG-----EFENDAPGIDNPAYNPKLTPPDQEEYLLINAV	431	
QY	317	GGTNGEFLPDGCLNNGD-----PALQKPSNGSDWYNDLMARKEFDARGMKMTWDEGDNN	371	
Db	432	GGVNYF-----GDLITTPA--KPSNDS--PTASKDFWSPNTWYPTWN--GEBA	476	


```

OY 75 PTLADINPOTGAPFGTDEMYNGVLVWAMYGACTNTDNNCGYRTGAAGDIPAMSARV 134
DB 484 A-----MNDKSKFFPD-----NKYAOYS-----SKRI 506
OY 136 RTFOKYSFTHGRVVVHAKMPYGDMLPAILMILPEDWYVYGNPRSGEIDIETIGNRDEKN 194
DB 507 NTOKKLSLKAGRVDFRAKLPFGDGVWPAIMLIPKDSYVGTMAASEIDVMEARGRLPGSV 566
OY 195 TGEFLGIQKMGSTWMBGPGDNDKRYLTSLPKHSDNN-----YGDNFHTFMDSPN 248
DB 567 SG-----TIFHGGQVPVNO-----SSGDDYHFEGOTFANDYHYVYWEED 608
OY 249 GLRFVDE-----NQALLDVPPLIDANPWPVDFWEMGKFWLPQYENDNPMAGCTNLA 302
DB 609 NIKRYVYDGKFFKYKYNQ-----QWYSTAEPNNPN-----A 638
OY 303 PEDONFPHILVAVGCT--NGFIPDGCINRGDPPALQKFPWSNGDWYNDAMRKFPDARGNW 360
DB 639 PFDEPFILMILVAVGNGDFGRTP---NASDIPA----- 669
OY 361 KMTWDEGDNNAMQVDYIRYK 382
DB 670 -----TMQVDYIRYK 680

RESULT 5
S48201
Licheninase (EC 3.2.1.73) precursor - Rhodothermus marinus
N:Alternate names: beta-glucanase; endo-beta-1,3-1,4-D-glucanase
C:Species: Rhodothermus marinus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
C/Accession: S48201
R:Spiliact, R.; Hregyldsson, G.O.; Kristjansson, J.K.; Eggertsson, G.; Palsdottir, A.
Eur. J. Biochem. 224, 923-930, 1994
A:Title: Cloning and sequencing of a Rhodothermus marinus gene, bglA, coding for a therma
A:Reference number: S48201; MUID:95010084
A:Accession: S48201
A:Molecule type: DNA
A:Residues: 1-286 <SPI>
A:Cross-references: EMBL:U04836; NID:9450182; PIDN:AAA60459.1; PID:9450183
A:Experimental source: strain 21
C:Genetics:
A:Gene: bglA
C:Keywords: glycosidase; hydrolase
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-286/Product: licheninase #status predicted <MAT>

-- Query Match 15.4%; Score 344.5; DB 2; Length 286;
-- Local Similarity 27.5%; Pred. No. 1.4e-20;
-- Matches 112; Conservative 37; Mismatches 96; Indels 163; Gaps 17;

OY 1 MRMWLVVCLIFGEFAPTD-----WDQYHIVODEDFY---DGAKHOEVATGGG 50
DB 13 MRRITFLSVLIGCSMLSDNSDKAPHE---LVMSDEDFYGLDPPEKMDYDVGCHGNG 69
OY 51 NSEFOLYODG--ANSFVRDGLFIRPTLADINIPOTGAPFGTDEMYNGVLVWAMYGAC 109
DB 70 NOELQYTRARERENARVGGVLIIEA---RHEPEYREY----- 105
OY 110 TMTDNGGCTRGAGADIPPMASARVTFQKYSTHGRVVVHAKMPYGDMLPAILMILPED 169
DB 106 -----TSARLVTRGKASMTYGRFIRRLRPSGRGTWPAIMLIPR 145
OY 170 WYVG--WPRSGEIDIETIG--NRD-----FKNTGGEFLGIQKMGSTWMBGPGMDNRY 220
DB 146 QYGSATYPPDNGEIDIMEVGNFNPVYVGTVHTKAYNHLGTRGGS----- 192
OY 221 WLTLSPKSHDDNNGYDNGFTFMDWSPNGIRFVDE-----NQALLDVPPLIDANPM 274
DB 193 --IRVPLARTD-----FHVYAIEMVPEELRMFVDSLVYRPNRNLTD---PEADWRHM 241
OY 275 WVDWFEMGKFWLPQYENDNPMAGCTNLAPEDFNPHILVAVGCTNGFIPDGCINRGDP 334

```

```

DB 242 -----PFDQPFILINIAVAGAMG-----OQGVDP 267
OY 335 ALQKFPWSNGDWYNDAMRKFPDARGMKWTWDEGDNNAMQVDYIRYK 382
DB 268 -----EAFPAQ-----LVVDYVRYR 283

RESULT 6
JC6150
Gram-negative bacteria-binding protein precursor - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 21-Jul-2000
C/Accession: JC6150; PC6039
R:Lee, W.J.; Lee, J.D.; Kravchenko, V.V.; Ulevitch, R.J.; Brey, P.T.
Proc. Natl. Acad. Sci. U.S.A. 93, 7888-7893, 1996
A:Title: Purification and molecular cloning of an inducible gram-negative bacteria-bi
A:Reference number: JC6150; MUID:96353914
A:Accession: JC6150
A:Molecule type: mRNA
A:Residues: 1-467 <LEE1>
A:Cross-references: GB:L38591; NID:91209252; PIDN:AAH40946.1; PID:91209253
A:Accession: PC6039
A:Molecule type: protein
A:Residues: 42-53; 56-83; 88-100 <LEE2>
A:Experimental source: fat body
C:Comment: This is a soluble protein having affinity for the Gram-negative bacterial
C:Keywords: glycoprotein; immune response
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-467/Product: gram-negative bacteria-binding protein #status predicted <MAT>
F:119/182/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 14.0%; Score 314; DB 2; Length 467;
-- Local Similarity 27.3%; Pred. No. 7.5e-18;
-- Matches 101; Conservative 58; Mismatches 143; Indels 68; Gaps 17;

OY 27 WMEDEFDYDGAQM--HEVTATGGNSEPOLYTDGANSFVRDGLFIKFTLADINIPQ 85
DB 147 IFEEGFDSDENWVQLEQYIPIYHPEYPFVSYQRNMLTVSTADGNLHI-----NAKLO 199
OY 86 TGAP--RGTFPMYNGVLVWAMYGACTNTDNNCGYRTGAAGDI--PPMASARVTFQKYSFT 143
DB 200 QHNPGLDLSITISGLTN---LFSGCTSS--AEACIKASGADILPEPVSGRI--TSIGFAFT 254
OY 144 HGRVVVHAKMPYGDMLPAILMILPEDWYVYGNPRSGEIDIETIGN-----RDFKNT 195
DB 255 YGTVEIRAKLPQGDWLYPELILPEFLAKTGSMTYASGYVYIACARGNAELVSGPNDSNT 314
OY 196 GGEFLGIQKMGSTWMBGPGMD--DNRYWLTSLPKHSDDNNGYDNGFTFMDWSPNGLEFFV 254
DB 315 -----VLXGFPIMDLCEBRENFLSTRKRRDGTSGDSFFHYYSVQMTDFIALSV 362
OY 255 DDENQALLDVPPLIDANPWPVDFWEMGKFWLPQYENDNPMAGCTNLAPEDFNPHILNY 314
DB 363 DGEEMARVAPRDAL-----PAYAAHPRHLLQASQAPDDHFTITLGV 408
OY 315 AVGTGNGFIPDGCINRGDPPALQKFPWSNGDWYNDAMRK---FEDAGNNKWTWDEGDN 370
DB 409 AAGGITFEF--RDGSIISGG---VTKPW-----RDSARKASVHWRHNSDMFPRNSQP--- 455
OY 371 NAMQVDYIRY 380
DB 456 --SLIVDFKV 464

RESULT 7
B72428
Laminarinase - Thermotoga maritima (strain MS8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C/Accession: B72428

```


Db 137 GTFPFDGDEMNITD--PTGFGVYGLDRNSSAKGLISANSNVYMAADSKHNSGGRPS 194

QY 134 VRFPQKSTFHHGVV---VYAKMPVGDWMLPAMLPEDWYGMFPGSGEIDIIETIGNR 190

Db 195 IRLQSTQYEHGHGFLIDLIH--LPYGGTWPAPFMTLGGD-----WPNGEIDIEGV--- 244

QY 191 DFKNTGGEFLGIQKMGSTMMHGPC--WDDNRWMLTSLPKHSDDDP----- 232

Db 245 ---NVG-----TSNOVTLHTGDCCEMEDIKRYMTGALTQNCVADPNSYNACGVENP 295

QY 233 ---NYGDNH-----TFWEDMSPNGLR--FFVDENQALLDVPYPLIDANPWVDFWENG 282

Db 296 SCPSYGEAFKRNKGGVFLDMRSEGRSMF-----NRS--EIPEDITSGSPQPA---KWS 346

QY 283 KP 284

Db 347 EP 348

LT 13

A:5132

Hypothetical protein PAB1790 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A75132

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: A75132

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1144 <KAM>

A:Cross-references: GB:AJ248285; GB:AL096836; NID:95458067; PIDN:CAB49770.1; PID:ej51566

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1790

Query Match 5.6%; Score 125.5; DB 2; Length 1144;
Best Local Similarity 18.4%; Pred. No. 0.038;
Matches 85; Conservative 49; Mismatches 132; Indels 197; Gaps 23;

QY 15 GFAFTMDQYHIWODEFDYF---DGAQMO-----HEVATGGGNS----- 52

Db 66 GLOGANLDKLYIAMDEYLYIAIKNTNANKVAYGIGIDVPGEGFSGESGDSWGKIN 125

QY 53 -----EFOLYTODGANSFVNDGKLFKPTLLADNINPOTGAPFEGTDFMYNGVLDMVAMYG 107

Db 126 FTREIDYEVY-----FWMGDKIM-----SDAFNMN-----CTDWEKXSISIEGEX- 167

QY 108 ACTNDNNGCYRGAAGDIPRAPSARVTFQKYSFTHGRVVAHAKMPVGMVLMPLMLP 167

Db 168 AYTGDSSSGIQ-----VLEVAIP-----NSALGKGP 193

QY 168 E-----DWYVGMFPRSGEIDIIETIGNRDKFNGEFLGIQKMGSTMMHMGCPMD-----N 218

Db 194 EKTAITTWI-AGEEGSSAVDTLPVDPISNINGE-----WTDADMLS 236

QY 219 RYMLTSLPK-----HSDWMN-----YGNFHTFWMFDSMNGLRFVDDENQAL 261

Db 237 NFIEISITKVIYDGNLDMNKAELVAGVPSGISEBANLDRLVSDSNLYIAITNTNTAK 296

QY 262 LDVYPL-IDAN-----PW-----WVDEWEGKP 284

Db 297 YMPDYGAIVDNGSGIGTYDPAKKIYSGTYLPDIYIAEAQDALTWVGCKW--- 353

QY 285 WLPQYENDNFWAGTINLAPDQNHFLINAVAGTNGFIPOGICINRGDPAALQPMNSGD 344

Db 354 -----DGNEMITGTNINSVGEYAY-----IGDNDKI-----QTEIKVPMWS--- 389

QY 345 WYDAMKRFDPARGNMTWTWDEGDNNAM-----QVDYIRV 380

Db 390 -----ALGMPERKAIIAMVAGNEDGNSAVDTLPVDPISIDYENI 428

RESULT 14

B69798

hypothetical protein yeta - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: B69798

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M

koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Laidino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; M01D:98044033

A:Accession: B69798

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-857 <KUN>

A:Cross-references: GB:Y99107; GB:AL009126; NID:92632866; PIDN:CAB12528.1; PID:g26330

A:Experimental source: strain 168

C:Genetics:

A:Gene: yeta

A:Superfamily: Bacillus subtilis hypothetical protein yeta

Query Match 5.6%; Score 125; DB 2; Length 857;
Best Local Similarity 19.8%; Pred. No. 0.029;
Matches 97; Conservative 53; Mismatches 162; Indels 178; Gaps 27;

QY 12 FEGFAFTMDQYHIWODEFDYFDGAK-----W----- 40

Db 289 FAHQTONAVMDVKIV--QDSSDHYSLSKRTGKDYAWGMLHGRKAGCYAGKNGVAL 347

QY 41 -----QHEVYATAGGNSERQ--LYTQDG-----ANSEFVRGKLF 73

Db 348 GLRYFEKYPALBITITGLAGSRPKRTIWLMPDDEANDLRHYTGNTFVAASAYEGFDEMRS 407

QY 74 KPTLLADNINQGTGAPFG--TDFMYNGVLDMVMA-----MYGACTWTDNNGCYRT 120

Db 408 DPTGIA-NTNEISLACFSNHPDSDEVNLNADKQAPPLICEPBY-----YES 455

QY 121 GAAG-----DIPRAPSARVTFQK-----YSFTHGRVVAHAKMPV----- 155

Db 456 KALGVWSIIDTSHPLKKELEQDLDAFLFYKKEVQQRWYGFHMGVDMHTYDPIRHMWR 515

QY 156 ---GDMLPAILMPLPDMVYVGMFPRSGEIDIIETIG--NRPFKNTG--GEFLGIQKMG 206

Db 516 YDLGYSAMONNELVPTTILMLOAFPRSGREDIFRMAEATRTSTSETSFHLGEYAGLSRH 575

QY 207 STMHNGPMDNRYMLTSLPKHSDWMN--GDNFHTFWMFDSMNGLRFVDDENQALL----- 262

Db 576 NVVHMGCGCKEAKRISMAGL-----HFFYLYTG-----DDRGLDLETVK 615

QY 263 DVYPLIDANPWVDFWENGKPMVPOYENDNP--WAGGTN--LAPD--QNFHILINAVG 317

Db 616 DADVALVKTDPRA--FEYKGRH--PTHARTGPDMNAFCSNMLAEWERFENSEYTKIEFG 672

QY 318 GTNGFIPOGICINR-----GG-----DPALOKPMSNGWYDAMARKF--PDARGNW----- 360

Db 673 -----INCLRLPLRLLSGPTREYDPATSMLHMGGIAGGYHMIITAFGAPQYMELA 725

QY 361 ---KMTWDD 366

Db 726 ELDDWENED 735

RESULT 15

A83748

endo-beta-N-acetylglucosaminidase BH0785 [Imported] - *Bacillus halodurans* (strain C-125)

```
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #text change 31-Dec-2000
```

C:\Accession: A83748
C:\Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C/Accession: A03/48
R: Takami, H.: Nakasone, K.: Takaki, Y.: Maeno, G.: Sasaki, R.: Masui, N.: Fuji, F.: Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A;Reference number: A83650; MUID:20263314

A;Accession: A83748

A;Status: preliminary

A;Molecule type: DNA
A;Positives: 1-878 <ST

A;Cross-references: GB:
A;ResIdues: 1-8/8 <SIU>

A: Experimental source:

C:\Genetics:

ene: BH0785

Query Match	5.28;	Score 116.5;	DB 2;	Length 878;
-------------	-------	--------------	-------	-------------

Best Local Similarity 22.18; Pred. No. 0.15;
Matches 88; Conservative 50; Mismatches 143; Indels 117

Matches	88;	conservative	50;	mismatches	143;	indels	11/;	gaps	29/
---------	-----	--------------	-----	------------	------	--------	------	------	-----

3 LVVCLLEGGFAIDWDQH-----LVWQDEFDYFDGAKW-----QHEVTAIGGN 31

```

Db      14 VLITLPSOGFASOPESSYWPETLLDWSPETD--PDARENRRSSITPLREVEY----- 64
      ::|||::|||::|||::|||

```

0V 52 SEFOLYTODGANSFVRDCKLEIKPTLLADNINPOT-GAPE--GTDENYNGVLDVWAMYGA 108

DATE	DESCRIPTION	AMOUNT	BALANCE
12/1	TO BALANCE		100.00
12/2	BY CASH	50.00	50.00
12/3	TO CASH	25.00	75.00
12/4	BY CASH	10.00	65.00
12/5	TO CASH	30.00	95.00
12/6	BY CASH	15.00	80.00
12/7	TO CASH	20.00	100.00
12/8	BY CASH	10.00	90.00
12/9	TO CASH	10.00	100.00
12/10	BY CASH	5.00	95.00
12/11	TO CASH	5.00	100.00
12/12	BY CASH	5.00	95.00
12/13	TO CASH	5.00	100.00
12/14	BY CASH	5.00	95.00
12/15	TO CASH	5.00	100.00
12/16	BY CASH	5.00	95.00
12/17	TO CASH	5.00	100.00
12/18	BY CASH	5.00	95.00
12/19	TO CASH	5.00	100.00
12/20	BY CASH	5.00	95.00
12/21	TO CASH	5.00	100.00
12/22	BY CASH	5.00	95.00
12/23	TO CASH	5.00	100.00
12/24	BY CASH	5.00	95.00
12/25	TO CASH	5.00	100.00
12/26	BY CASH	5.00	95.00
12/27	TO CASH	5.00	100.00
12/28	BY CASH	5.00	95.00
12/29	TO CASH	5.00	100.00
12/30	BY CASH	5.00	95.00
12/31	TO CASH	5.00	100.00

[illegible]

QY 109 CTNTDNNGCYRTGAGD--IPPAMSARVRTFQKYSFTHGRVVHAK--MPV-GDWLWP 161

```
Db 113 MV-----YWAGSAGEGITTPTS-----GDVIDAHRNGVPILGNVFFP 150
```

.....

[illegible]

Db 151 -----PK-----EWDKMLVRGEDGSFPADKLLVAEYYGFDGWFINQ 195

QY 220 YWLTSLPKHSDW-----NYGDNFHTFWD--WSPNGIRF--FVDDENQALLDVPY 266

106 EEEEEEETAVWEEETVIOEURECUCUUTMTWDCATCUCDIPBONWY EEDENAEEEOACN 35E

CY 067 BTDDNDEETMEATEWNCW DUT POWEIMDIEA CCENT A REDONEUETT MIVVCCNGCEI - - - 333

26/ PLIDANPWWVDIEWENGK PWLPQYENDBNPWAGGTLNAPEDQNEHFLNVAVGGINGT--- 323

Search completed: October 11, 2001, 15:55:28
Job time: 286 sec

Job time: 286 sec


```

Db 424 NLIMDEFGNTLDTSTKMYETGYLYLNNDPATWGMNAELQHYNTSTQNYVYDQGLNIK 483
OY 75 PTLADNINPQATPGFTDPMNGVLVDMVAMYGACTITDNGCYRTAGADIPPMASARY 134
Db 484 A-----MDSKSFPODP-----NRTAQS-----SGKI 506
OY 135 RFQKYSFTHGRRVYVAKKMPGDMPLMPLMPLMPLMPLMPLMPLMPLMPLMPLMPLMPL 194
Db 507 NTRDKSLATYGRDPAKLEPTGCGVAPALMMLTKDSYGYGMAASGELDVAKRGLPGSV 566
OY 195 TGGEEFGLGCKMSTGMNGPCMDNRYWLSLPHSDWN-----YGDNFTFTWFDWSPN 248
Db 567 SG-----THFGGQMPVNO-----SSGCDYHPPEGQTFANDHYVSVVEED 608
OY 249 GLREFYVDE-----NOALLDVPRPLDANPMWVDPEWEGCKPMLPOYENDNPMAGCTNLA 302
Db 609 NIKWYVDGKFFYKVTNO-----QWYTAAPNPN-----A 638
OY 303 PEDQNFHTLNAVAGT-----NGEIPDCINRGDPALOKPMSNGDWYNDAMRKFFDARGNW 360
Db 639 PDEPEYLLMNLAVGNGFDGGRTP-----NASDLPA----- 669
OY 361 KTWDEGDNNAQVDYIRYK 382
Db 670 -----TMQVYVRYK 680

RESULT 2
GUB_RHOMR
ID GUB_RHOMR STANDARD: PRT: 286 AA.
AC P45798;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENDO-BETA-1,3-1,4 GLUCANASE)
DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE).
GN BGLA.
OS Rhodothermus marinus.
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.
OX NCBI_TaxID=29549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21 / rti-378;
RC MEDLINE=95010084; PubMed=7925416;
RA Spilliaert R., Hreggvidsson G.O., Kristjansson J.K.,
RA Eggertsson G., Palsdottir A.,
RA "Cloning and sequencing of a Rhodothermus marinus gene, bglA, coding
for a thermostable beta-glucanase and its expression in Escherichia
coli."
RL Eur. J. Biochem. 224:923-930(1994).
CC -1- FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN
CC BUT NOT ON CMC CELLULOSE OR XILAN. THE ENZYME HAS A TEMPERATURE
CC OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC
CC LINKAGES IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U04836; AAA60459.1; -
CC DR HSSP: P23904; IAKK.
CC DR InterPro: IPR000757; -
CC DR Pfam: PF00722; Glyco_hydro_16; 1.
CC DR PROSITE: PS01034; GLYCOSYL_HDROL_F16; 1.
CC DR Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 286 BETA-GLUCANASE.

```

```

FT ACT SITE 158 158 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 163 163 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 286 AA; 33145 MM; 7215C3624135191 CRC64;

Query Match 15.4%; Score 344.5; DB 1; Length 286;
Best Local Similarity 27.5%; Pred. No. 2e-21;
Matches 112; Conservative 37; Mismatches 96; Indels 163; Gaps 17;

OY 1 MRWTLVLCLEFGEFAFTD-----WDQYHYWQDEFYF-----DGAKWQHEVTANGG 50
Db 13 MRRATFLSLYLIGCSMLGSDRSRKAPHE---LWMSDEFYSGLPDEKMDYDVGSHGMG 69
OY 51 NSEFQLTQDG-ANSFVRDGLFKPILLADNINPQATPGFTDPMNGVLVDMVAMYGAC 109
Db 70 NQELQYTRARIENARVGGGLITEA-----RHEPYBGRY----- 105
OY 110 TNDNNGCYFTGAGDIPPMASARVRFQKYSFTHGRRVYVAKKMPGDMPLMPLMPLMPL 169
Db 106 -----TSARLVTRGRKASWYTGREIRARLPSCGCTWPAIMLPDR 145
OY 170 WYVGG--WPRSGEIDILETG-NRD-----FKMTGEEFLGCKMSTGMNGPCMDNRY 220
Db 146 QYGSAYWPDNGEIDIMEHGFNDVYHGVTRKAYNHLGTORGS----- 192
OY 221 WLTSLPKHSDDMNKGDNFHTFWMFSPNGLRFYVDE-----NOALLDVPRPLDANPMW 274
Db 193 --INVTARTD-----FHYAIEWTPEIEWFDSDSLYFPERLRD---PEADMKHW 241
OY 275 WDFEWEGKFPMLPOYENDNPMAGCTNLAPEPDNFTLNAVAGTNGEIPDCINRGDP 334
Db 242 -----PFDQPHLLMNLAVGAMSG-----QQGVDP 267
OY 335 ALQKPSNGDWYNDAMRKFFDARGNKKMTWDBEDNNAQVDYIRYK 382
Db 268 -----EAPPAQ-----LVVDYVRYR 283

RESULT 3
EXSH_RHIME
ID EXSH_RHIME STANDARD: PRT: 465 AA.
AC Q33680;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ENDO-1,3-1,4-BETA-GLUCANASE EXSH (EC 3.2.1.-) (SUCCINOGLYCAN
DE BIOSYNTHESIS PROTEIN EXSH).
GN EXSH.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteri; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC YORK G.M., Walker G.C.;
RA "The Rhizobium meliloti exoX gene and prsD/prsE/exsH genes encode
RT components of independent degradative pathways which contribute to
RT production of low-molecular-weight succinoglycan."
RL Mol. Microbiol. 25:117-134(1997).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98226741; PubMed=9560202;
RA YORK G.M., Walker G.C.;
RA "The Rhizobium meliloti ExoX and ExsH glycanases specifically
RT depolymerize nascent succinoglycan chains."
RL Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS. PERHAPS BEFORE
CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC AGGREGATION STATE.

```

CC -1- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D89164; AAB64093.1; -
 CC InterPro: IPR000757; -
 CC InterPro: IPR001343; -
 CC Pfam: PF00722; Glyco_hydro_16; 1.
 CC PROSITE: PS00353; hemolysin_cabind; 1.
 CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 CC PROSITE: PS00330; HEMOLYSIN_CALCIDIUM; FALSE_NEG.
 CC EXPOLYSACCHARIDE synthesis; Glycosidase; Hydrolase; Plasmid.
 CC DOMAIN 275 465 NUCLEOPHILE (BY SIMILARITY).
 CC ACT_SITE 349 349 CATALYTIC.
 CC ACT_SITE 354 354 PROMOT DONOR (BY SIMILARITY).
 CC SEQUENCE 465 AA; 50286 MW; 6C8482366B9B8CA8 CRC64;
 SO

Query Match 7.8%; Score 175.5; DB 1; Length 465;
 Best Local Similarity 21.6%; Pred. No. 2.8e-07;
 Matches 77; Conservative 40; Mismatches 102; Indels 137; Gaps 15;

QY 27 VQDEPFYFGAKWQHEVTATGGNSEFQLYTQDANSFVADGKLFKPTLLADINPQT 86
 DB 238 VMAKFI-----VMAPEKGTILSSNGPDQWYINPS-----VEPTA-----SVNPF 277
 QY 87 GADPGDFMTNGVLDVAMATGA-CTNNDNNGCYRTGAGADIPAMSAVRFQKSPFHG 145
 DB 278 -----VANGVLTITRAAPASEAIOAEINQYDT-----SGMLTYSFPAQYIG 319
 QY 146 RVVYHAKMPVGDMLPAMLMPEDMVYGGWPRSGEIDIIETIGNRDFKNTGEGFLGLOKM 205
 DB 320 YFEKRAMDPDQGVWPAFWLLPAD---GSRP---PELDVVEKRGD---SNT-----V 363
 QY 206 GSRMHNGPGDNDKRYWLTSLPKHSDDMNYGDNFTFMDSPNGIRFVVDENQALLDVP 265
 DB 364 IATVH--SNETGSTSIENSVKAD---ASGFHTYGVLTMEETIYWFDDAIAIARADTP 417
 QY 266 YPLIDANPMWVDWEMGKPMPLPQYENDNPMWAGTINLAFDQNFHILVAVGTFNGTIPD 325
 DB 418 SDMDH-----PMY-----MLVNLAVGAGTAGTPRD 441
 QY 326 GCINRGDPAALQKFWNSGDMWYNDAMRRFFDARGNMKWTWDEGDNNAQVDYIRYV 381
 DB 442 GL-----ADGSEKKIDYIKAY 457

RESULT 4
 EGIC_RHIME STANDARD; PRT; 465 AA.
 AC 092302;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ENDO-1,3-1,4-BETA-GLUCANASE EGIC (EC 3.2.1.-) (SUCCINOGLYCAN
 DE BIOSYNTHESIS PROTEIN EGIC).
 GN EGIC.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OC NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CXM1-105;
 RX MEDLINE=99413305; PubMed=10485295;
 RA Sharypova L.A., Yurel S.N., Keller M., Simarov B.V., Puehler A.,
 RA Becker A.;
 RT "The eft-482 locus of Sinorhizobium meliloti CXM1-105 that influences
 RT symbiotic effectiveness consists of three genes encoding an
 RT endoglycanase, a transcriptional regulator and an adenylate cyclase.";
 RL Mol. Gen. Genet. 261:1032-1044(1999)
 CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
 CC SUCCINOGLYCAN. DYNAMICALLY REGULATES THE MOLECULAR WEIGHT
 CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
 CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
 CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
 CC AGGREGATION STATE (BY SIMILARITY).
 CC -1- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ225896; CAB38101.1; -
 CC InterPro: IPR000757; -
 CC InterPro: IPR001343; -
 CC Pfam: PF00722; Glyco_hydro_16; 1.
 CC Pfam: PF00353; hemolysin_cabind; 1.
 CC PRINTS: PR00313; CABINDGRPT.
 CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 CC PROSITE: PS00330; HEMOLYSIN_CALCIDIUM; FALSE_NEG.
 CC EXPOLYSACCHARIDE synthesis; Glycosidase; Hydrolase; Plasmid.
 CC DOMAIN 275 465 NUCLEOPHILE (BY SIMILARITY).
 CC ACT_SITE 349 349 CATALYTIC.
 CC ACT_SITE 354 354 PROMOT DONOR (BY SIMILARITY).
 CC SEQUENCE 465 AA; 49602 MW; 1BC4A640685B9654 CRC64;
 SO

Query Match 7.5%; Score 167; DB 1; Length 465;
 Best Local Similarity 21.4%; Pred. No. 1.4e-06;
 Matches 84; Conservative 38; Mismatches 113; Indels 158; Gaps 18;

QY 16 FAFTMDQYHI-----VQDEPFYF---DGAK-----WQHEVTATGGG 50
 DB 196 FANTSAADOLHANOFSLADRSVLTQTFSDDFYTLQSDSTSGVMPKYMMAPEKATLLG 255
 QY 51 NSEFQLYTQDANSFVRDQKLEIKPTLLADINPQTGA-PFGTDFMYNGVLDVAMATGA- 108
 DB 256 NDELQWY-----VNPT-----YQPASANPFS---VTDGLVLTITAKPASQ 292
 QY 109 CTNNDNNGCYRTGAGADIPAMSAVRFQKSPFHGRVYVAKMPVGDMLPAMLMPE 168
 DB 293 AIQAEFTNGYDYLT-----SGMLTYSFPAQYIGFEMRADMPDQDAGAPAFWLLPG 342
 QY 169 DMVYGGWPRSGEIDIIETIGNRDFKNTGEGFLGLOKMSTMGSPWMDNRYWLTSLPKH 228
 DB 343 D---GTWP---PELDVVEKRGD---PNTVYATVHNETGQTS-----IASAAY 384
 QY 229 SDDMNYGDNFTFMDSPNGIRFVVDENQALLDVPYPLIDANPMWVDWEMGKPMPLQ 288
 DB 385 TDT-----SGFHYGVLTMEETIYWFDDAIAIARADTPSDMDH-----PM 424
 QY 289 YENDPMWAGTINLAFDQNFHILVAVGTFNGTIPDGCINRGDPAALQKFWNSGDMWYND 348
 DB 425 Y-----MLVNLAVGAGTAGTPDGLM---GG----- 446
 QY 349 AMRKFEDARGNMKWTWDEGDNNAQVDYIRYV 381
 DB 447 -----AEKQVDYIKAY 457

```

RESULT 5
DHET_ACEPO STANDARD; PRT; 738 AA.
ID DHET_ACEPO
AC P28036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
GN ADHA.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NB11028;
MEDLINE=91159482; PubMed=2001402;
Tamaoki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
*Cloning and sequencing of the gene cluster encoding two subunits of
membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.*;
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
REDUCED ACCEPTOR.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL, D00635; BAA00528.1; -
DR PIR, S14270; S14270.
DR HSP, P22619; 2MTA.
DR
DR InterPro: IPR000345; -
DR InterPro: IPR001479; -
DR Pfam: PF01011; Bacterial_POO_1; 1.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C_1; 1.
KW Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;
SO

```

Query Match 5.0%; Score 112.5; DB 1; Length 738;
 Best Local Similarity 20.4%; Pred. No. 0.072;
 Matches 78; Conservative 37; Mismatches 113; Indels 155; Gaps 21;

```

OY 48 GAGNSEFOLYTDGANSFVR-----DGKLFKPTLLADININPOTGAPFGND-FMYN 97
DB 212 GNGSGSE-----GAGPFSAPDAEKGKVDWRFYTP-----NRKNEDDASDSVLAN 258
OY 98 GVLVDWAMYGACTNTDNGC-----YRTGAGD----- 125
DB 259 KAYQWSPFGAMTRGCGGGVWMDSIYDPAVLVYLGVGNGSPWNKRYRESGGDNFLG 318

```

```

OY 126 -----IPPAMGARVRFQ-----KSTFTHR-----VYVHAK----- 152
DB 319 SIYALKRGESELYVWHFETPMDDWFTSDQIMTLIDLPINGETRHYIYHARKNGFFYIID 378
OY 153 MPVGDMLPMAIWMPEWVYGGW-----PRSGEIDIEITGNRD--FKATGGEFLGID-K 204
DB 379 AKTGEFI-----SGKRYVYVMAASGLDPKIG-----RPIYNPDALYLTGKEWIGIFGD 427
OY 205 MG-----STMHWPGMDNDRRWLTSLP-----KHSDDWNYGDNFHTFWEFD 244
DB 428 LGHNFAAMAFSP-----KIGLYIIPAQOVPELYTNQVGFTHPPDSSWNLGLDMKVKGIP 482
OY 245 WSPVGLTEFFYDDENQALLDVPYPLIDANPMWYDWE---NGKPHLPQYENDNPPAGGTNL 301
DB 483 DSPPAKQAFVYKDLK-----GVIYAMDPOKQAEAW--RVDHGGPMNGGILA 525
OY 302 APFDONFETILN--VAVGTING 321
DB 526 TGGDLRFQGLANGEFHAYDAING 548

```

RESULT 6

```

ID XYNL_RUMFL STANDARD; PRT; 954 AA.
AC P29126;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BIFUNCTIONAL ENDO-1,4-BETA-XYLANASE XYLA PRECURSOR (EC 3.2.1.8).
GN XYNL.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Ruminococcus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17;
RX MEDLINE=92261318; PubMed=1584021;
RA Zhang J.-X., Flint H.J.;
RT "A bifunctional xylanase encoded by the xynA gene of the rumen
RT cellulolytic bacterium Ruminococcus flavefaciens 17 comprises two
RT dissimilar domains linked by an asparagine/glutamine-rich sequence.";
RL Mol. Microbiol. 6:1013-1023(1992).
CC -1- FUNCTION: XYLANASE DOMAIN 1 RELEASES MORE XYLO-OLIGOSACCHARIDES
AND DOMAIN 2 MORE XYLOSE.
CC -1- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF 1,4-BETA-D-XYLOSIDIC
LINKAGES IN XYLANS.
CC -1- PATHWAY: XYLAN DEGRADATION.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
F (FAMILY 10 OF GLYCOSYL HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL, Z11127; CAAT7476.1; -
DR PIR, S18043; S18043.
DR PIR, S20907; S20907.
DR HSP, P48793; 1XND.
DR InterPro: IPR001000; -
DR InterPro: IPR001137; -
DR Pfam: PF000331; Glyco_hydro_10; 1.
DR Pfam: PF00457; Glyco_hydro_11; 1.
DR PRINTS: PR00134; GLHYDRLASE10.
DR PRINTS: PR00911; GLHYDRLASE11.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; 1.

```


[illegible]

```

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RA MEDLINE=95066320; PubMed=7922030;
RX Crenells S., Garman E.F., Laver W.G., Vimer E.R., Taylor G.L.;
RT "Crystal structure of Vibrio cholerae neuraminidase reveals dual
RT lectin-like domains in addition to the catalytic domain.";
RL Structure 2:535-544(1994).
CC -!- FUNCTION: CLEAVES THE TERMINAL SIALIC ACID (N-ACETYL NEURAMINIC
CC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS PROVIDING FREE
CC SIALIDS WHICH CAN BE USED AS CARBON AND ENERGY SOURCES.
CC SIALIDS HAVE BEEN SUGGESTED TO BE PATHOGENIC FACTORS IN
CC MICROBIAL INFECTIONS. NANP FACILITATES CHOLERA TOXIN BINDING
CC TO HOST INTESTINAL EPITHELIAL CELLS BY CONVERTING CELL SURFACE
CC POLYSIALOGLYCOSIDES TO GM1 MONOGLYCOSIDES.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC
CC LINKAGES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEURAMINYLS
CC RESIDUES TO GALACTOSE, N-ACETYLBEXOSAMINE, OR N- OR O-ACETYLATED
CC NEURAMINYL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS,
CC GLYCOLIPIDS OR COLLOIDAL ACID.
CC -!- COPROCTOR: CALCIUM.
CC -!- SUBUNIT: MONOMER (PROBABLY).
CC -!- INDUCTION: MAY BE CONTROLLED BY SIALIC ACID AVAILABILITY AND A
CC GROWTH-PHASE-DEPENDENT MECHANISM.
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -!- SIMILARITY: CONTAINS 4 BNR REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; M83562; AAA27546.1; -.
DR EMBL; AE004255; AAP94933.1; ALT_INIT.
DR EMBL; M19268; AAA27547.1; -.
DR PIR; A43866; A43866.
DR TIGR; VC11784; -.
DR PDB; 1KIT; 05-JUN-97.
DR InterPro; IPR002860; -.
DR Pfam; PF02012; BNR. 4.
KM Hydrolase; Glycosidase; Signal; Repeat; Calcium; 3D-structure.
KT SIGNAL 1 24
FT CHAIN 25 781 STALIDASE.
FT REPEAT 263 274 BNR 1.
FT REPEAT 585 596 BNR 2.
FT REPEAT 653 664 BNR 3.
FT REPEAT 718 729 BNR 4.
FT ACT_SITE 618 618 PROTON DONOR (BY SIMILARITY).
FT CONFLICT 22 22 S->L (IN REF. 3).
SQ SEQUENCE 781 AA; 85609 MW; FA85ED907FB2DAF0 CRC64;

Query Match 4.9%, Score 109; DB 1; Length 781;
Best Local Similarity 23.6%; Pred. No. 0.15;
Matches 82; Conservative 31; Mismatches 107; Indels 128; Gaps 20;

44 VTATGGNSG---POLYNODGAN---SEVRQSKLFIRKPTLLADININPQTGAFFGDPMYN 97
||| | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 143 VATATTAETHKFLVELVPLSPNSPSAFPEDK-----LRINDTP-----TASKON 189

98 GVLDWVMYAGACTNNDNNGCYRT---GAAGDPPAMSARVFVEOKYSFTGHVVVAHKAP 154
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 MI--VMG---NGSNNDYGAAIARDIKFELQGDVIYFGPRPIBSIVASVTYPGVYFAEKR 245

155 VGDMWLPAIMLPEDMVGAMP--RSGEIDIETIGNBDPKNTGGEEFLGIQMGSTMHG 212
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 246 VS-----GDPGALSNTNDII-FTRSD-----GTW- 272

213 PGWDNRKRVLISLPHASDDNMGDNFHFWPDMSNGLARFPVDENQALLDVPLYLD-- 270
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 273 ---DTENLTAEQIVNSDEPFSD-----PRDIYDS 300

```


QY 271 ANPMWVDEWENG-----KPLDPO--YENDNPWAGTINLAPD-----ONEHFI 311
 DB 301 SNTVYVYARMPETDAQNDRIKPMWPNNGIFYSYDVASGNQAPLDYDQKERSEFQI- 359
 QY 312 LNVAVGTNGEIPDCINRGDPAQKPMWSNGDWYNDARKRFEDAGN 359
 DB 360 --AGWGSELYRRTSLN-----SQDQWQSNKIRIYDGAAN 394
 RESULT 9
 KRE6_YEAST STANDARD: PRT: 720 AA.
 AC P32486;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-GLUCAN SYNTHESIS-ASSOCIATED PROTEIN KRE6 (KILLER TOXIN-
 RESISTANCE PROTEIN 6).
 GN KRE6 OR YPR159M.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OX Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 RX MEDLINE-92107936; PubMed-1837148;
 RA Roemer T., Bussey H.;
 RT "Least beta-glucan synthesis: KRE6 encodes a predicted type II
 membrane protein required for glucan synthesis in vivo and for glucan
 synthase activity in vitro.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:11295-11299(1991).
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95176711; PubMed-7871892;
 RA Roemer T.D., Fortin N., Bussey H.;
 RT "DNA sequence analysis of a 10.4 kbp region on the right arm of yeast
 chromosome XVI positions GPH1 and SGV1 adjacent to KRE6, and
 identifies two novel tRNA genes.";
 RL Yeast 10:1527-1530(1994).
 RN CHARACTERIZATION.
 RP MEDLINE-95014745; PubMed-7929594;
 RA Roemer T., Paraviciuni G., Payton M.A., Bussey H.;
 RT "Characterization of the yeast (1-->6)-beta-glucan biosynthetic
 components, Kre6p and Skn1p, and genetic interactions between the
 PKC1 pathway and extracellular matrix assembly.";
 RL J. Cell Biol. 127:567-579(1994).
 CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF (1->6)- AND (1->3)-BETA-D-
 GLUCAN POLYMERS OF THE YEAST CELL WALL. IN VIVO. IT IS REQUIRED
 FOR FULL ACTIVITY OF BETA-GLUCAN SYNTHASE IN VITRO. IT MAY BE A
 BETA-GLUCAN SYNTHASE, PART OF A MULTIPROTEIN GLUCAN SYNTHASE OR
 A MODULATOR.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI
 SUBCOMPARTMENT.
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- SIMILARITY: STRONG, TO SKN1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@sdb-sdb.ch).
 CC EMBL: M80657; AAA34726.1; -
 DR EMBL: L33835; AAB59312.1; -
 DR PIR: A41624; A41624.
 DR SGO: S0006363; KRE6.
 KW Glycoprotein; transmembrane; Cell wall; Signal-anchor;
 FT Phosphorylation.
 FT DOMAIN 1 252 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 253 273 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT DOMAIN 274 720 (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 538 538 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 563 563 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 720 AA; 80178 MW; 958989EF7E9ACD69 CRC64;
 Query Match 4.8%; Score 108; DB 1; Length 720;
 Best Local Similarity 19.0%; Pred. No. 0.16;
 Matches 82; Conservative 48; Mismatches 124; Indels 178; Gaps 23;
 QY 36 DGAKMOHEVYATAGGNGSEFQOLYODGANSFVRDQKPIKFTLLADINPOTGAPFGTDM 95
 DB 322 DSKW-----ELVFSDEFNA---EGRTFY-----DGDPPYWTAP---DVH 355
 QY 96 YNGVLDV-WAMTGACTNTD-----NNGCYRTGAAGDIPRAMSAVRFQKYSFT 143
 DB 356 YDATTQDLEWYSPDASTTNGTQLRLRDARKNHLY-----YRSGMLQSMNKVCF 405
 QY 144 HGRVYVHAKP-VG--DWLPAIWLPE-----DMVYGGMPRS----- 178
 DB 406 QGALFISANLPNRYGRVSGLPGLMTWGNLGRPEYLASTQGVWYSYSCDAGITPNQSSP 465
 QY 179 -----GETDIETIGNRDPKNTGCEFFLGIQMGST 208
 DB 466 DGISYLPQGLKSLCTCDVEDHPNQGVRGAPETDVE--GETPTK-----ICVGIASQS 517
 QY 209 MHMGPEMDNRWLTSLPRKHSDDMNTGDNFHTFWMSPRGLEFFEDENQALLDVPYPL 268
 DB 518 LQIAP-----FDIMWPDYDFEIVYFTTTMTATYAGRPQOAVSAVSTL 562
 QY 269 IDANPMWVDEWENG--KPLDPOYENDNPWAGTINLAPEDONFHTLNVAVGTNGFI-- 323
 DB 563 ---NWYWEFEGEGYEQYKAYALEYND-----DDNGY--IRWEVGTPTFTYIH 605
 QY 324 ----PDGCTINRGDPAQKPMWS-----NGCW-YNDARKRFEDAGNKKMTWDEGDN 370
 DB 606 AKALHPDG--NIGWRRISEPMISILNLGISNMWAYIDMOYIEFPV----- 649
 QY 371 NAMQVDYIRYVK 382
 DB 650 -VMSIDYVRIYQ 660
 RESULT 10
 GUB_PAEPO STANDARD: PRT: 238 AA.
 AC P45797;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENDO-BETA-1,3-1,4 GLUCANASE)
 GN (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE).
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Paenibacillus.
 OX NCBI_TaxID=1406;
 RN SEQUENCE FROM N.A.
 RP STRAIN=ATCC 842;
 RX MEDLINE-92041687; PubMed-1938968;
 RA Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
 RT "Two beta-glucanase genes are clustered in Bacillus polymyxa:
 molecular cloning, expression, and sequence analysis of genes
 encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
 RL J. Bacteriol. 173:7705-7710(1991).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC LINKAGES
 IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.

CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X57094; CAA40379.1; -
CC HSP: p23904; ICPM.
CC InterPro: IPR000757; -
CC Pfam: PF00722; Glyco_hydro.16; 1.
CC PRINTS: PR00737; GLHYDRLASE16.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC Hydrolyase; Glycosidase; Signal.
CC SIGNAL 1 26
CC CHAIN 1 27 238 BETA-GLUCANASE.
FT ACT_SITE 129 129 NOCLOPHILE (BY SIMILARITY).
FT ACT_SITE 133 133 PROTON DONOR (BY SIMILARITY).
FT DISULFID 56 85 BY SIMILARITY.
SQ SEQUENCE 238 AA; 26919 MW; C0CE7BAEASD40B8C CRC64;

Query Match 4.7%; Score 106; DB 1; Length 238;
Best Local Similarity 22.0%; Pred. No. 0.07;
Matches 62; Conservative 32; Mismatches 92; Indels 96; Gaps 16;

QY 25 HIWODEPFYDCAKMOHETATGSGNSEFQLYTQDGAN-SYVRDGLFIKPLTLADNIN 83
DB 28 NFWF-EPILSFNSSTWQ---KADGYSNGQWFCNCTWRANNNFINDGLKLTSSPANN-- 81
QY 84 POTGAFGTDFMNGYLDVWAMTACTNTDNGCYRTGAAGDIPPMASARVTFQKSF- 142
DB 82 -----KFDC-----GEVRSNTNNG-YGLYEVSMKPAKNTGVSSFFYTGP 121
QY 143 THGRVYVNAKMPYGDMLPRAIMLPEDWYGGMPRSGEIDIIETIGNRDKRTNGEFLGT 202
DB 122 SHG-----TQW-----DEIDI-----EFLG- 136
QY 203 OKKSGTMHMGDMNRWYLTSLPKHSDDMNG--DNHTEFMDSPMGLRFVVD-- 256
DB 137 -KDTTVQF-----NTITNGVGHKEKITNLGPDASTSFHTAFDQPGITIKMYDGVLT 188
257 ENQALDVPPLIDANPWWYDFEWG---KPLPQYENDNP 294
DB 189 KHATATNIP-----STPGKIMMLMNGTGVDSMLGSGVNGANP 225

RESULT 11
DHET_ACEEU STANDARD; PRT; 739 AA.
ID 044002: 007952;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, last sequence update)
DE 15-JUL-1999 (Rel. 38, last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
GN ADH.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconacetobacter.
OX NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DES11 / DSM 6160;
RA Thurner C.A.K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
CC REDUCED ACCEPTOR.
CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).

CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X82894; CAA58066.1; -
CC EMBL: Y09480; CAA70688.1; -
CC InterPro: IPR000345; -
CC DR InterPro: IPR001479; -
CC DR InterPro: IPR002372; -
CC Pfam: PF01011; Bacterial_POQ_6.
CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC DR PROSITE: PS00190; CYTOCHROME_C; 1.
CC Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT CHAIN 1 35
FT BINDING 36 739 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BINDING 651 651 HEME (COVALENT) (BY SIMILARITY).
FT METAL 654 654 HEME (COVALENT) (BY SIMILARITY).
FT METAL 655 655 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;

Query Match 4.7%; Score 106; DB 1; Length 739;
Best Local Similarity 20.6%; Pred. No. 0.25;
Matches 78; Conservative 37; Mismatches 118; Indels 146; Gaps 20;

QY 48 GGGSEFQLYTQDGANSEFVR-----DGKLEIKPLTLADNINP----- 84
DB 212 GNGSGSE-----GAKFTVAFAETAETGKVDWREFTAP-----NPKNEPHITASDSYLM 258
QY 85 -----QTGAPPGT-----DFWNGYLDVWAMTACTNTDNGCYRTGAAGD-- 125
DB 259 NKAQTMTSPGAMTRGGGGGVWDSIYDPAVDL-VYLGVNGSGPMWVKYRSEKGNLFF 317
QY 126 -----TPPMASARVTFQ-----KYSTHGRVYVNAKMPYGDMLPRAIMLP-- 167
DB 318 LGSIVALKPETGEYVWHFOETPMQMDFTSVQIIMTLDLPINGETRHVIAHPKNGFEYI 377
QY 168 -----EDWYVGW-----PRSGEIDIIETIGNRD--FKNTGGEFLGQ-KMG-- 206
DB 378 IDAKTGEFISGKNIVYVNMASGLDPKGTG-----RPITNPDLALYLLTGKEWGIQEDDLGGH 432
QY 207 --STMHNGPGWDMNRWYLTSLP-----KHSDDWNYGDNFTFEDWSPN 248
DB 433 NFAAMAFSP-----KTGLVYIPAAQVPELYTNGVGFPHRDSNMLGLDMNKVGIPODPE 487
QY 249 GLRFFVDDENQALDVPPLIDANPWWYDFE---WGKPLPQYENDNPAGGTLAFLAP 305
DB 488 AKQAFVVDLK-----GWLAVMDPKQKQAEAW--RVDHKRPMWNGGILATGSD 530
QY 306 QNEHFIILN---VAVGNG 321
DB 531 LILFGLANGEHAYDATING 549

RESULT 12
OSTA_HAEIN STANDARD; PRT; 782 AA.
ID 044846;
AC P44846;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)

DE ORGANIC SOLVENT TOLERANCE PROTEIN HOMOLOG PRECURSOR.
 GN IMP OR OSTA OR H10730.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae.
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips R., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd".
 SC Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, TO THE E. COLI HOMOLOG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U3756; AAC22389.1; -
 DR TIGR: H10730; -
 DR Outer membrane; signal.
 FT SIGNAL 1 23
 FT CHAIN 24 782
 FT POTENTIAL.
 FT ORGANIC SOLVENT TOLERANCE PROTEIN
 FT HOMOLOG.
 FT SEQUENCE 782 AA; 90084 MW; E73C8A5786B02D1B CRC64;
 SQ

Query Match 4.7%; Score:106; DB 1; Length 782;
 Best Local Similarity 21.3%; Pred. No. 0.26; Mismatches 112; Indels 148; Gaps 26;
 Matches 83; Conservative 46; Mismatches 112; Indels 148; Gaps 26;

32 FDYFDGAKMOHEVYATGGNSEFQLYTODG--ANS--FV-RDGKLFKPTLLADNINPQ 85
 106 FDYKD-----NQIMML-GKDAEFNLSHDGMLTSEYEFVGRGR-----GKADNIT-- 151
 86 TGAPFGTDFMYNGVLDVYVAMTGACTNDNNGCYRTGAAGDIPPAASA-RVRTFOKYSFT- 143
 152 -----LHNN-----TRVKNMNTFTSLH-----GDVAMVADASEIRQYEEYAE 191
 144 --HGRVVVHA-----KAPVG-----DWLM-----PAIMML----- 166
 192 MHAHREIKHGVYPTTLPDLPIDGRNRSGLLISAGTSSODGLWYAOPIYINWLNIPDTL 251
 167 ---PEDWYVGMPRSGEIDLETIGNDFKNTGGEFLGIQKSGTMHNGPGWDMRMYLT 223
 252 TETPKYMSRGMQANGERRVLTSGE--GKVAGEYLG-----KVRSSEVASDR----- 298
 224 SLPHASDMMVGNDFHFTFPMSPNGLRFVYDDENQALLDVYPLIDANPMWVDFWE-WG 282
 299 --KRHLFYWMNNSFLQMW-----RLNINVTFSVSKRYFENDEDSIYG 338
 283 KP---WLPYENDMPWAGTMLAPFDONFHLINAVAGTGNGF--IPDCINNGGPA--- 335
 339 RSTDTGANOT-----ARATYQPNYNSLS-----AHQFQFDIDIVNIGRYAVAPQ 384
 336 -----LQKPMNSGDMYNDAMKRPFD 355
 385 LDFNYHKYDLANGMLNFKLHSAVR--FD 411

RESULT 13
 XYND_RUMFL STANDARD; PRT; 802 AA.
 ID XYND_RUMFL
 AC 05317;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLANASE/BETA-GLUCANASE PRECURSOR (INCLUDES: ENDO-1,4-BETA-XYLANASE
 DE (EC 3.2.1.8) (XYLANASE); ENDO-BETA-1,3-1,4-GLUCANASE) (EC 3.2.1.73)
 DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE)).
 GN XYND.
 OS Ruminococcus flavefaciens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Ruminococcus.
 OX NCBI_Taxid=1265;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-17;
 RX MEDLINE=93259938; Pubmed=8491715;
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;
 RT "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
 RT glucanase domains, encoded by the xynd gene of Ruminococcus
 RT flavefaciens".
 RT J. Bacteriol. 175:2943-2951(1993)
 CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
 CC BETA-1,3-1,4-GLUCANASE ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
 CC LINKAGES IN XYLANS.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC LINKAGES
 CC IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.
 CC -1- PATHWAY: XILAN DEGRADATION.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY
 CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
 CC GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: S61204; AAB26620.1; -
 DR HSSP: P23904; IATK.
 DR InterPro: IPR000757; -
 DR InterPro: IPR001137; -
 DR Pfam: PF02018; CBD_6; 1.
 DR Pfam: PF00457; Glyco_hydro_11; 1.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRASE16.
 DR PRINTS: PR00911; GLHYDRASE11.
 DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11_1; 1.
 DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_2; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR Xylan degradation; Hydrolase; Glycosidase; signal;
 KM Multifunctional enzyme.
 FT SIGNAL 1 31
 FT CHAIN 32 802
 FT DOMAIN 32 244
 FT DOMAIN 245 523
 FT DOMAIN 524 555
 FT DOMAIN 556 802
 FT ACT_SITE 124 124
 FT ACT_SITE 226 226
 FT ACT_SITE 684 684
 FT ACT_SITE 524 529
 FT DOMAIN 532 543
 FT DOMAIN 546 553
 FT POLY-THR.
 FT POLY-THR.
 FT POLY-THR.
 FT SEQUENCE 802 AA; 89091 MW; 2880A689647284AF CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:56:43 ; Search time 34.86 Seconds
(without alignments)
1457.405 Million cell updates/sec

Title: US-09-596-101a-3
Perfect score: 2240
Sequence: 1 MRMTLVVLCILFGESEFAFTD.....DDEGNNMAYDIIVYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Maximum number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	100.0	384	5 077072	077072 eisenia foe
2	716.5	32.0	361	5 090064	090064 pacifastacu
3	713	31.8	489	5 026660	026660 strongyloce
4	703	31.4	395	5 017492	017492 anopheles g
5	390.5	17.4	495	5 09N189	09N189 bombyx mori
6	369.5	16.5	877	2 045095	045095 bacillus ci
7	357	15.9	492	5 09VVR5	09VVR5 drosophila
8	353	15.8	1324	2 059328	059328 clostridium
9	350	15.6	276	2 052754	052754 rhodotermu
10	350	15.6	435	2 068641	068641 oerskovia x
11	329.5	14.7	487	5 09NJ98	09NJ98 manduca sex
12	323	14.4	494	5 09NHO	09NHO drosophila
13	314	14.0	467	5 017233	017233 bombyx mori
14	307	13.7	306	2 051333	051333 oerskovia x
15	302	13.5	481	5 096363	096363 hyphanta
16	295	13.2	490	5 09VSR4	09VSR4 drosophila
17	295	13.2	490	5 09NHA8	09NHA8 drosophila
18	287	12.8	442	2 060039	060039 thermotoga
19	286.5	12.8	422	2 092690	092690 flavobacter

20	284	12.7	642	2 09WXN1	09WXN1 thermotoga
21	275.5	12.3	285	2 09F3A0	09F3A0 streptomyces
22	257.5	11.5	297	1 073951	073951 pyrococcus
23	249.5	11.1	410	5 09VVR4	09VVR4 drosophila
24	249.5	11.1	461	5 09NHA9	09NHA9 drosophila
25	239	10.7	422	2 09ENR5	09ENR5 streptomyces
26	209	9.3	673	5 027082	027082 tachypleus
27	190.5	8.5	478	2 09L816	09L816 streptomyces
28	184.5	8.2	845	2 09KMF3	09KMF3 clostridium
29	179.5	8.0	466	2 088021	088021 streptomyc
30	148	6.6	364	14 084415	084415 parametium
31	138.5	6.2	294	2 007242	007242 mycobacteri
32	132	5.9	419	3 09USW3	09USW3 schizosacch
33	125.5	5.6	1144	1 09VOD4	09VOD4 bacillus su
34	125	5.6	857	2 031530	031530 pyrococcus
35	121.5	5.4	424	3 060019	060019 pacifia rho
36	116.5	5.2	878	2 09KER4	09KER4 bacillus ha
37	115	5.1	284	10 P93670	P93670 hordeum vul
38	111	5.0	282	2 09MOD2	09MOD2 arabidopsis
39	110.5	4.9	290	2 09R119	09R119 aeromonas s
40	110.5	4.9	856	3 074170	074170 aspergillus
41	108	4.8	282	10 080803	080803 arabidopsis
42	108	4.8	636	3 013789	013789 schizosacch
43	108	4.8	720	3 006472	006472 saccharomyc
44	107.5	4.8	545	2 084907	084907 zebellia ga
45	107.5	4.8	725	3 09P670	09P670 neurospora

ALIGNMENTS

RESULT 1
077072 PRELIMINARY; PRT; 384 AA.
ID 077072;
AC 077072;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE COELOMIC CYTOLYTIC FACTOR 1.
GN CCF1.
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
CX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98406152; PubMed=9733802;
RA Beschin A., Biles L., Hanssens F., Raymakers J., Van Dyck E.,
RA Reyers H., Brys J., Gomez J., De Baetselier P., Timmermans M.,
RT "Identification and cloning of a glucan- and lipopolysaccharide-
binding protein from Eisenia foetida earthworm involved in the
activation of prophenoloxidase cascade."
RT J. Biol. Chem. 273:24948-24954(1998).
RL EMBL: AF030028; AAC35887.1;
DR J. EMBL: AF030028; AAC35887.1;
SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

Query Match 100.0%; Score 2240; DB 5; Length 384;
Best Local Similarity 100.0%; Pred. No. 1e-182;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRMTLVVLCILFGESEFAFTDMDQYHIVWODEDFYDGAKKMOHEVATGSGNSEFOLYTD 60
DB 1 MRMTLVVLCILFGESEFAFTDMDQYHIVWODEDFYDGAKKMOHEVATGSGNSEFOLYTD 60
QY 61 GANSEVRDGLKFIKFTLLADNINPOTGAPFGTFYXNGVLVWAMYGACTNDNNGCYRT 120
DB 61 GANSEVRDGLKFIKFTLLADNINPOTGAPFGTFYXNGVLVWAMYGACTNDNNGCYRT 120
QY 121 GAAGDIPPMASARVTFQKYSFTHGRRVVVHAKMPYGDMLPAILMILPEDWVYGGMPRSGE 180
DB 121 GAAGDIPPMASARVTFQKYSFTHGRRVVVHAKMPYGDMLPAILMILPEDWVYGGMPRSGE 180

Query Match	Best Local Similarity	31.4%	Score 703;	DB 5;	Length 395;
Matches 146;	Conservative	58;	Mismatches 121;	Indels 38;	Gaps 11;
26	IVMODEFDYFDGAKMOHEVTATGCGNSEPOLYTODGANSFVBDGKLFKPTLLADNINPQ	85			
62	LIFEDNFEDEDEKMEHVNTLAGGCMWERYWTNNRNSFEVDGALNIRPTLLADQ----	117			
86	TPAPGTPDMYNGVLDVWAMYG--CTNFDNNGCTRTGAAGDI-PPAMARVITOKISF	142			
118	----FGLDFMTGTLSTLQSYPTDHCRTNDATFYGCVRVGNROHIVPVKSAIRITISSEF	173			
143	TGRRVVAHAKMVGMDLMPAIWMLPDMYVGGMPRSGEIDITETIGNRFFKNTNGGFLGI	202			
174	KYGRAEVRAKLTGTGDMLEWPAIMLLPRKRAYGIVWPASGENDLHESKRENNLYLDGYO-IGT	232			
203	QKMGSTMHMGPCWMDNRVYMLTSLPKHS-DMWYGDNFHTFWEDWSPNGLFFVDDENQAL	261			
233	RGVGTGLHGFNPSPYNGYPTATLTKNALPEGEFSKSFYFGEVWMPEDNITVSINGEDLAT	292			
262	LDPVPLDIANDWVDFEWMGKPLPQYENDNPWAGITLAPFDONFHITLVANVGITNG	321			
293	IG-----GDFWTRG--GFDKHLIENDPWRHGTMAEPDEFPHITILAVAGV-A	337			
322	FIPDGINGGPALOKPMSNGDWMYNDARKFFDAGNMKMWMD--DEGDNNAMQVDTI	378			
338	FEPDAATNFGG-----KPKNNNS--PQAAITDWMGRHAQMLPTMNLERDGGKSAISLVDIV	390			
379	RVY 381				
391	KVM 393				
RESULT 5					
ID 09NL89	PRELIMINARY;	PRT;	495	AA.	
AC 09NL89;					
DT 01-OCT-2000 (TREMBLRel. 15, Created)					
DT 01-OCT-2000 (TREMBLRel. 15, Last sequence update)					
DE 01-OCT-2000 (TREMBLRel. 15, Last annotation update)					
DE BETA-1,3-GLUCAN RECOGNITION PROTEIN PRECURSOR.					
OC Bombyx mori (Silk moth).					
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylidae;					
OC Bombycoidea; Bombycidae; Bombyx.					
OC NCBI_Taxid=7091;					

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KINSHU X SHOWA; TISSUE-HEMOCYTES;
RX MEDLINE=20138243; PubMed=10671539;
RA Ochiai M., Ashida M.;
RT "A pattern-recognition protein for beta-1,3-glucan. The binding domain
RT and the cDNA cloning of beta-1,3-glucan recognition protein from the
RT silkworm, bombyx mori."
RL J. Biol. Chem. 275:4995-5002(2000).
DR EMBL; AB026441; BAA92243.1; -
KW Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 495 BETA-1,3-GLUCAN RECOGNITION PROTEIN.
SQ SEQUENCE 495 AA; 55802 MM; 092490397456BF2 CRC64;

Query Match 17.4%; Score 390.5; DB 5; Length 495;
Best Local Similarity 30.6%; Pred. No. 3e-25;
Matches 114; Conservative 57; Mismatches 122; Indels 79; Gaps 19;

QY 26 IWMODEFD--YFDGAKWQHEVATGGSNEFQLYTQDGSANFVRDGLKLFKPTLLADNIN 83
:::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 185 LTFEDQFNPIPIHRGKINWPEVAFKPEBPDPFENVYLSD--NAEYDNGKILIRKATLE---- 238
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 84 PGTGAPFGTDFMYNGVLDMVAMYAGCTINTDNG--CYRTGAAGD--IPPMASRVFTQKY 140
::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 239 ---SKYGEDFVRQS-LD---LSEKCTGVGTACQURE-ASGPLILPPIITAKISTRHQF 289
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 141 STHGRVYVHAAMPYGDWIMPAIMWLPBEMVYGGMP-RSGEIDIIETIGNRPFKNTGGEF 199
::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 290 ARKYGVETRAAMPKGMWLYPELLEPRNITGVARNYASGLIKIASYKGNAEF----- 342
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 LGIOKMGSTMHMGCP---GWDNRMYLTSLPKHSDDMYGDNFHTFWDSPGLRFFVDD 256
::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 343 ---SKLTVAGPIMTGSDPYRSYLIKENITEESNN--NDPHNTLEMRPGIITLLVDG 394
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 257 EN-----QALLDVPYLIDANPMWVDFEWGKFWLPQYENDNPAGSTNLAPDQNFH 309
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 395 SEYGEIKPEGEFYNY-----ANSYKVEA-----APQ-----WLKGTIMAPDELFY 435
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 310 FLINVAVGSTNGETFDGCGINRGSDPALOKPWSNGSDMYNDAMKRFDPARGNMKWTMDDED 369
::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 436 VSTGLNVAGIRFESD-----ISNRPKNS--ATKALMLKFDANSQWEPYDWD-- 482
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 370 NNAMQVDYIRYV 381
::|::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 483 -SALOVDYKYVF 493

RESULT 6
ID 045095 PRELIMINARY; PRT: 877 AA.
AC 045095: 053369;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE BETA-1,3-GLUCANASE BGLH PRECURSOR.
GN BGLH.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IAH165;
RX MEDLINE=94033858; PubMed=7764221;
RA Yamamoto M., Aono R., Horikoshi K.;
RT "Structure of the 87-kDa beta-1,3-glucanase gene of Bacillus circulans
RT IAH165 and properties of the enzyme accumulated in the periplasm of
RT Escherichia coli carrying the gene."
RL Biosci. Biotechnol. Biochem. 57:1518-1525(1993).
DR EMBL; D17519; AAC60449.1; -
EMBL; S67033; BAC60453.1; -

```


DE ENDO-1,3(4)-BETA-GLUCANASE (EC 3.2.1.6) (ENDO-1,4-BETA-GLUCANASE)
DB (ENDO-1,3-BETA-GLUCANASE) (LAMINARIINASE).
GN LICA.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
CC Clostridium.
RN NCBI_Taxid=1515;
RP [1]
RC SEQUENCE FROM N.A.
RA STRAIN=DSM 1337;
RL Schwarz W.H., Schilling S., Fuchs K.P., Staudenbauer W.L.;
CC Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOPOLYESTERS OF 1,3- OR 1,4-LINKAGES IN
CC BETA-D-GLUCANS WHEN THE GLUCOSE RESIDUE WHOSE REDUCING GROUP IS
CC INVOLVED IN THE LINKAGE TO BE HYDROLYSED IS ITSELF SUBSTITUTED AT
CC C-3.
DR EMBL; X89732; CAA61884.1; .
DR HSSP; P14090; IULO.
DR InterPro; IPRO001119; .
DR InterPro; IPRO03305; .
DR Pfam; PF00395; SLH; 3.
DR Pfam; PF02018; CBD_6; 4.
DR Hydrolase; Glycosidase.
SQ SEQUENCE 1324 AA; 148202 MW; 03BE40FE9E09538AB CRC64;

Query Match 15.8%; Score 353; DB 2; Length 1324;
Best Local Similarity 27.2%; Pred. No.1.6e-21.
Matches 100; Conservative 51; Mismatches 93; Indels 124; Gaps 16

OY 22 DOYIHYMODEEP--YFDGAKWQHEVTANGSGSERPOLYTDQGANSFVRDKLFKEPTLLA 79
DB ::::| | | | : : : : : : | | | | : : : : | : : :
OY 425 EEMRWMSDEFNGSEIMNMAINSYDDPTNGRMNGEVOSTYQN--NAVIKDGALVTEAR--- 479
DB ::::| | | | : : : : : : | | | | : : : : | : : :
OY 80 DNINQGTAGPEPTDFMYNGVLDVNMATGCCTTDDNNCGCYRTGAAGDIIPRMAGARVTFQK 139
DB 480 ---KEDIPESETGYHT-----SSKLTKCK 503
OY 140 YSFTHGRVVNAKMPGVGDWLPAIMLPAPEDW--VGGMPSGEIDITETIGNRDNFKTNGE 198
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 504 KSMWKGKEIRAKMKOGGQISIPAIMMPADEPFYGWPKGEDIMELLGHER----- 556
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 199 FLGIQKMGSTMMWGPGMDNRNYWLSLPKHSDDMWNTGDNFHFWFWDSPNGLRFPYVDEN 258
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 557 ---DKIGTGIHFGEPRHKESGGTYT-LP---EGQTFADDPFYHSIEWERGETIRWTIDGR- 607
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 259 QALLDYPPPLIDANPMWVDWFEMGKFWLPQYENDNPWAGG--TNLAPFOQHFTILNAV 316
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 608 --LVHV-----ANOW-----YRRD--PYLADDVITYPAFPQNFELLINISIV 644
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 317 GGTCNGFIPEGCINRGDDPALCPKWSNGDWYNAMRKFFPARGNMKMTWDEGDNNAQYD 376
DB 645 G-----GWMFG-----YPETTYF-----PQQAVYD 665
OY 377 YIRVYKRKN 384
DB ::::| | | | : : : : | | | | : : : : | : : :
OY 666 YVRYVKQKD 673

RESULT 9
AC 052754 PRELIMINARY; PRT; 276 AA.
ID 052754
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE LAMINARIINASE.
GN LAMR.
OS Rhodothermus marinus.
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.
OX NCBI_Taxid=29549;
RN [1]
RP SEQUENCE FROM N.A.

Query Match	15.6%	Score 350;	DB 2;	Length 276;
Best Local Similarity	28.5%	Pred. No. 4e-22;		
Matches 116;	Conservative 39;	Mismatches 92;	Indels 160;	Gaps 19;
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries R., Krah M.:			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF047003; AAC69707.1; -			
DR	InterPro; IPR000757; -			
DR	Pfam; PF00722; Glyco hydro.16; 1.			
DR	PROSITE; PS01034; GLYCOSYL-HYDROL.F16; 1.			
DR	SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;			
RC	STRAIN-IT1278:			
RA	Borries			


```

0Y 85 QTAPAFGEFDE--MYNGVLDLWMAWYAGACTINTD--NGCY----RFGAADDIPMASARRTG 137
Db 228 -----YRRODLSIANSRLD---LSEKCTGTHRRICEJLHSTGSPSPIMPIYPRSTK 279
0Y 138 OKYSETHGRVVVYHAAKMPGVMDLMPALMMLP-EDWVYVGMPRSGEIDITIEFGNRDEKNTG 196
Db 280 ETEAFOYRIETIRAKLPRGMDIYVLLLEPLETEYOGSGYESGOLRALALRGNSVLMR 359
0Y 197 GEFILGIOMKGSMTMHGPMQMDNRYMLSLPKHSDW-----NYGDNFTHFEDWSPNG 249
Db 340 GKLYD-----GRSLYGR-----VLSDAHOREDLMLSKRISHFGDFHTYLSIDMSSNR 389
0Y 250 LRFVYDDENOALLDYPPYLIDANPMWVDFEWEMGRPMLPYEENDPMAGTGNLAPFONFH 309
Db 390 LLEFVYDGVYGEMLNGFTELDEND-----RMOGGPMAPFDEMF 429
0Y 310 FILNVAVGCTGTFIPDGCINNGGPPALOKPWSNGDWTNDAMRKEFFDARGWKW--TWDE 367
Db 430 ISLGVSVSGFGEFDE-----VDHLRTATEYEKWPAT-----IIPRSCSSYRADOWMLPTWKOP 480
0Y 368 GDNNAWMOVDYIRVY 381
Db 481 ---ALKIDYRVF 490

```

```

RESULT 13
ID 017233
AC 017233
PRELIMINARY; PRT; 467 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GRAM-NEGATIVE-BINDING PROTEIN PRECURSOR (GNBP).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 42-53; 56-83 AND 88-100.
RT TISSUE=Fat BODY, AND HEMOLymph;
RX MEDLINE=96353914; PubMed=8735372;
RA Lee W.J., Lee J.D., Kravchenko V.V., Ulevitch R.J., Brey P.T.;
RT "Purification and molecular cloning of an inducible gram-negative
RT bacteria-binding protein from the silkworm, Bombyx mori.";
RT Proc. Natl. Acad. Sci. U.S.A. 93:7888-7893(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE INSECT-IMMUNE PROTEIN. IT HAS
CC -1- STRONG AFFINITY TO THE CELL WALL OF GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: SECRETED IN HEMOLymph.
CC -1- TISSUE SPECIFICITY: Fat BODY AND TO A LESSER DEGREE IN
CC -1- CUTICULAR EPIDERMAL CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS 6 HOURS AFTER AN INJURY
CC -1- OR MICROBIAL CHALLENGE.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
CC -1- SIMILARITY: TO H.CRENA GRAM-NEGATIVE-BINDING PROTEIN.
CC -1- SIMILARITY: WEAK, TO FAMILY 16 OF GLYCOSYL HYDROLASES.
DR EMBL; L38591; AAB40946.1; -.
KW Signal; Glycoprotein; Insect immunity.
FT SIGNAL 1
FT CHAIN 18
FT CHAIN 19 467
FT CARBOHYD 119 119
FT CARBOHYD 182 182
FT SEQUENCE 467 AA; 52217 MW; 780652F6C89046B77 CRC64;
Query Match 14.0%; Score 314; DB 5; Length 467;
Best Local Similarity 27.3%; Pred. No. 8.9e-19;
Matches 101; Conservative 58; Mismatches 143; Indels 68; Gaps
QY 27 VQDEEDYFDGAKMO-HEVYTAGGENSEFQLYTDGANSFYRCKLTIKPTLLADINPQ 85
DB 147 IFEEDSDYDENWMDYEDYPIYHEHYEYFSVSRNNLTSTADNLMH-----NAKIQ 199

```

```

QY      86  TGAF-FGTDFPNYNLVDVWVMNYGACTWTDNNNGCYRTGCAADI--PPASASAVFRFEOKYSFT 143
Db      200  QHMFGFLDDSTISGTLN---LFSGCTSS-ADACTKQASGADILPEIYSGRI--TISGFAFT 254
QY      144  HGRVYVNAKMPVGMWMLPAMLPALMEDWVYGGWP--RSGEIDIILEIGN-----RDFKNT 195S
Db      255  YGYEIRAKRLPQGMWLYPELLLEPFLKKGYSMTASGVVATACARGAELYSGPNDY SNT 314S
QY      196  GGEFLGIQKMGSTHMGPGMD-DNRWYMLTSLPKHSDDMNYGDNFHTFWFMDSPNGILRFVY 254S
Db      315  -----VLYGGPIIMDLBERENFSLTKRRRDSGTSMGDSFHTYSVQWTFDEFIALSV 362S
QY      255  DDENQALLDVEYPLPIDANPMWVDFWEMGKPMWLPQYENDNPMAGCTNLAPFDQNFHTLVN 314S
Db      363  DGEEMARVEAPRDAL-----PAVCNAPRHLLQAGSQNAPEDDHFYITLGV 408S
QY      315  AVGTGNGFIPDGCINRGSDPALOKPWSNGDMYINDAMK-----FFDANGANKKWTWDEEDGN 370S
Db      409  AAGGITTEF-RDGSITISGG---VTKPW-----RDSAKASVHFHRHMSDFPRMSQP--- 455S
QY      371  NAMQDYIYR 380
Db      456  -SLIVDFVKV 464

```

[illegible]

```

Db 199 -----TVA-GGYSDDN--GIMGTQHPQMSFADDFHTFGIDWTPEETWLVLD--GQ 246
QY 260 ALDVEYPLIDANPWWVDWFEMGKFWLPQYENDNPAGTNLAPFDONFHEILNVAAGT 319
Db 247 EYHRTVTVADYGANQM-----VEDDPFLILNVAAGG- 277
QY 320 NGTIPDGCIRGGDPALQKPMNSGDMYNDAMRKFFDARGNKMTWDEGDNNAQVYIR 379
Db 278 -----QMPGNPDATPEP-----QOKKVDYVR 299
QY 380 VY 381
Db 300 VY 301

```

```

RESULT 15
096363 PRELIMINARY; PRT; 481 AA.
096363;

```

```

01-MAY-1999 (TREMblrel. 10, Created)
01-MAY-1999 (TREMblrel. 10, Last sequence update)
01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE GRAM-NEGATIVE-BINDING PROTEIN (FRAGMENT).
GN GNBp1.
OS Hyphantria cunea (Fall webworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuidae; Arctiidae; Hyphantria.
OX NCBI_TaxID=39466;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99035790; PubMed=9818384;
RA Shin S.W., Park S.-S., Park D.-S., Kim M.G., Kim S.C., Brey P.T.,
RA Park H.-Y.;
RT Isolation and characterization of immune-related genes from the fall
RT webworm, Hyphantria cunea, using PCR-based differential display and
RT subtractive cloning.
RL Insect Biochem. Mol. Biol. 28:827-837(1998).
CC -1- FUNCTION: MAY BE INVOLVED IN THE INSECT-IMMUNE PROTEIN. IT PRESENT
CC STRONG AFFINITY TO THE CELL WALL OF GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS A FEW MINUTES AFTER AN
CC INJURY OR MICROBIAL CHALLENGE THEN STAYS RELATIVELY HIGH FOR 12
CC OR 24 HOURS.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
CC -1- SIMILARITY: TO H.CUNEA GRAM-NEGATIVE-BINDING PROTEIN.
CC -1- SIMILARITY: WEAK, TO FAMILY 16 OF GLYCOSYL HYDROLASES.
EMBL: AF023916; AAD09290.1;
KW Glycoprotein; Insect Immunity.
FT NON_TER 1
FT CARBOHYD 110 110 N-LINKED (GLCNAC..)(POTENTIAL).
SQ SEQUENCE 481 AA; 53014 MW; 49B2AC8C406929F6 CRC64;

```

```

Query Match 13.5%; Score 302; DB 5; Length 481;
Best Local Similarity 28.6%; Pred. NO. 9.7e-18;
Matches 107; Conservative 51; Mismatches 146; Indels 70; Gaps 19;

```

```

QY 27 VMODEPFFGAKWQHE---VTATGGNSSEFOLYTQDANSFVADGKLFKPTLLADNI 82
Db 155 IFEENFTREDVWQIYIPVSTPEFVSQHLSDPTVA-VTGGNLRITPKL----- 208
QY 83 NPOTGAPFGTD-FMYNGLVDWAMYGACTNTDNNGCYRTG-AAGDIPPMASARYTFQKY 140
Db 209 --QORMGFTDSSYSSSLNT---FSGCT-APARACMKDAGASLTPVSGRI-TSKAF 261
QY 141 SETHGRVVAHAKAFVGMWLMFAIMLPEWVYGGNP-RSGEIDIIETIGNR-----DF 192
Db 262 AFYTGTVFAKAKLPQGMWIPPEILPEFLKRYGSTHYSSGVIKIASANGNRELTSGYTDY 321
QY 193 KNT---GGEPLGICQKMSSTMWGPQMDNRYMLSLPRHSDMNTGDNFHTFMDSPNG 249

```

```

Db 322 SNKMIFGSPVNMNQ-CYDTLLESKASSNGROW-----GDDFHEVYLMRAPER 367
QY 250 LRFVYDENQALLDVPYPLIDANPWWVDWFEMGKFWLPQYENDNP---WAGSTNLAPEDQ 306
Db 368 ITLSVDGVEWAVRVEPTASGLSGR-----FPQTSKLPRTFLAAGTKNAPEDD 414
QY 307 NEHFLNVAAGSTNGFTPDGCIINRGDPALQKPMNSGDMYNDAMRKFFDARGNKMTWMD 366
Db 415 HFYTLGLVAAGSTTEF-PDGVQTSGRP---KPMWNTG--SKAMLFHWMEDMSRPAITNQ 468
QY 367 EGDNNAMQVYIRV 380
Db 469 P----QLLVQYKV 478

```

```

Search completed: October 11, 2001, 15:56:45
Job time: 303 sec

```